ESSAYS ON U.S. ENERGY MARKETS

A Dissertation

by

DAVID A. BRIGHTWELL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2008

Major Subject: Economics

ESSAYS ON U.S. ENERGY MARKETS

A Dissertation

by

DAVID A. BRIGHTWELL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	John R. Moroney
Committee Members,	Richard K. Anderson
	Steven L. Puller
	Richard T. Woodward
Head of Department,	Lawrence Oliver

August 2008

Major Subject: Economics

ABSTRACT

Essays on U.S. Energy Markets. (August 2008)

David A. Brightwell, B.S., Illinois State University;

M.S. Illinois State University

Chair of Advisory Committee: Dr. John R Moroney

This dissertation examines three facets of U.S. energy use and policy. First, I examine the Gulf Coast petroleum refining industry to determine the structure of the industry. Using the duality between cost-minimization and production functions, I estimate the demand for labor to determine the underlying production function. The results indicate that refineries have become more capital intensive due to the relative price increase of labor. The industry has consolidated in response to higher labor costs and costs of environmental compliance.

Next, I examine oil production in the United States. An empirical model based on the theoretical framework of Pindyck is used to estimate production. This model differs from previous research by using state level data rather than national level data. The results indicate that the production elasticity with respect to reserves and the price elasticity of supply are both inelastic in the long run. The implication of these findings is that policies designed to increase domestic production through subsidies, tax breaks, or royalty reductions will likely provide little additional oil. We simulate production under three scenarios. In the most extreme scenario, prices double between 2005 and

2030 while reserves increase by 50%. Under this scenario, oil production in 2030 is approximately the same as the 2005 level.

The third essay estimates demand for fossil fuels in the U.S. and uses these estimates to forecast CO_2 emissions. The results indicate that there is almost no substitution from one fossil fuel to another and that all three fossil fuels are inelastic in the long run. Additionally, all three fuels respond differently to changes in GDP. The result of the differing elasticities with respect to GDP is that the energy mix has changed over time. The implication for forecasting CO_2 emissions is that models that cannot distinguish changes in the energy mix are not effective in forecasting CO_2 emissions.

DEDICATION

To my parents and grandmother

ACKNOWLEDGEMENTS

I would like to thank everyone who guided me to completion of this dissertation. Dr. Wayne Ault of Southwest Illinois College recommended that I take a microeconomics course to improve my understanding of public policy analysis. Dennis Shannon taught that course and introduced me to the wonderful field of economics and completely changed my educational path. Dr. Daniel Rich of Illinois State University provided much advice on courses and other preparation for graduate school and helped me in the decision to attend Texas A&M.

Dr. John Moroney, my committee chair, provided numerous evaluations of my research and always provided valuable feedback to help me improve the final product. In addition to all his time and effort with this dissertation, he also introduced me to the field of energy economics and showed me the importance of the subject to every facet of society. My other committee members: Drs. Richard Anderson, Steven Puller, and Richard Woodward also provided helpful guidance and support.

All the secretaries and staff in the department were very helpful in all my endeavors. However, Tyffanne Rowan and Teri Tenalio deserve special mention for all the assistance they provided.

Finally, I would like to thank my friends and family for all their support and assistance. My mom provided financial help and endless encouragement. Karthik Easwar and Josh McCallie provided several much needed distractions and were always fun.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Fossil fuel energy is vital to any developed economy. It is necessary for the production and transport of goods and services, the lighting and heating of homes and schools and to recreational activities too numerous to list. The importance of fossil energy to both the maintenance and growth of the economy poses many public policy topics regarding its usage. Fossil fuels now account for 88 percent of all commercial energy in the United States. Their importance will not soon diminish.

This dissertation examines three areas of public policy involving public management of domestic energy usage. Chapter II addresses the returns to scale and market structure of the refining industry. The increase in gas prices along with companies such as ExxonMobil reporting record profits have caused some members of Congress to question the competitiveness of the refining industry and whether federal regulation is needed in order to reduce prices for consumers.

This research finds that although the number of operable refineries in the United States has decreased from 301 to 149 between 1982 and 2003, the total volume of refined products declined by only 6 percent. The reduction in the number of plants was offset by increasing production capacity of existing plants and technological progress in

---

This dissertation follows the style of the *American Economic Review.*

the refining process. Overall, the refining industry substituted capital for labor, which became relatively more expensive, and increased plant size. Increased capital intensity in refining is motivated by current environmental regulation and increasing labor costs. There is no evidence that current high prices are the result of collusive behavior. Regulating the market price is likely to distort the allocation mechanism and create more government bureaucracy. Both of these effects are welfare decreasing.

Chapter III analyzes U.S. oil supply. President Bush, several Congressional leaders, and oil company executives have advocated reducing U.S. dependence on foreign sources of oil through increasing the domestic supply. Proposals include increasing access to government lands such as the Alaska National Wildlife Refuge and offshore areas, providing tax breaks for exploration and development, reducing the royalties to the government for oil pumped from federal lands, and providing subsidies to help offset the costs of pumping oil. We use state level data that includes prices, reserves, and barrels of oil produced between 1982 and 2004. These data account for over 98% of all domestic production. The models indicate that both the short run and long run price elasticities of supply are inlelastic. Policies such as royalty reductions or subsidies will provide little additional domestic oil. The output elasticities with respect to reserves are also inelastic. The implication is that further access to government lands or tax breaks for exploration and development will add only modestly to domestic oil supply.

We further test the model by comparing simulations of various pricing and reserve addition scenarios to Energy Information Administration (EIA) projections of domestic oil consumption between 2005 and 2030. The simulations show that with a 3% annual price increase and full development of the Alaska National Wildlife Refuge, production in 2030 will be slightly less than that in 2005. The EIA (2007) projects constant consumption growth of 2.5% annually. In this best case scenario in terms of total domestic production, oil prices will approximately double and reserves will increase by about 50% of the 2005 levels and the gap between domestic consumption and production still increases. These simulations clearly show that policies designed to stimulate production have little impact on reducing dependence on imported oil.

Chapter IV analyzes the domestic demand for fossil fuel energy. The chapter begins by estimating the demands of coal, oil, and natural gas. Estimates indicate that markets for these energies are highly separable. Nearly all coal is used to produce electricity. Natural gas and oil can be used for electricity generation but gas is most widely used for residential heating, and refined oil is most widely used for transportation. Neither coal nor natural gas is a feasible substitute for oil in transportation. Coal is not a feasible substitute for heating oils and the burners necessary to generate electricity from gas and oil are sufficiently different than those for coal that substantial capital investment is required to change the electricity generation mix.

After estimating energy demand, the chapter addresses the impact of fossil energy usage on carbon dioxide emissions. Our estimates show that coal produces about 90 million metric tons of CO_2 per quadrillion Btu, oil about 65.6 and natural gas about 51.5. These estimates are used to compare various forecasts of CO_2 emissions. Previous studies have treated energy as a composite good rather than separating each fuel into different markets. Forecasts based on separate markets are more accurate than results based on composite energy usage.

This result is counter to most forecasting outcomes. Traditionally, it is thought that parsimony leads to more precision. The key to our finding is that parsimonious models fail to account for the change in energy mix that occurred over time. A model that cannot distinguish the CO_2 content of coal from that of natural gas loses important information that is vital to providing an accurate forecast.

This dissertation shows the complexity of domestic energy markets. Public policies tend to focus on the supply side of markets because suppliers are easier to regulate than consumers. However, these problems require a balance of supply and demand incentives. The high current prices of gasoline and refined products are the result of increasing world demand for oil. Domestic demand has also increased faster than domestic supply. Designing policies that reduce domestic demand will lower market prices for gasoline. Additionally, these policies will reduce U.S. dependence on foreign oil and CO_2 emissions.

CHAPTER II

THE DEMAND FOR LABOR AND RETURNS TO SCALE IN PETROLEUM REFINING

Introduction

Industrial economists have struggled to answer the question of optimal plant size for decades. Stigler (1958) proposed analyzing the optimal size by using a survival approach. He suggested that plants that were too small or too big would eventually exit the industry because costs were too high to compete with plants of the optimal size. By tracking the patterns of plant size over time, the researcher could determine what sized plants performed best because plants of that size would prevail and more costly plants would cease operations.

Stigler's approach was useful given the data limitations of 1958. As data quality and availability improved, econometric estimation of cost functions or production functions began to emerge. Scherer (1980) reported that the minimum efficient scale of a refinery was 200,000 barrels and that there was a 4.8 percent increase in the average cost at one-third the minimum efficient scale. This indicates the long run average cost curve for crude oil refining is rather flat, suggesting economies of scale.

A more recent study by Hibdon and Mueller (1990) compared estimates of optimal refinery size based on Stigler's survival method to estimates derived from models of cost as a function of linear and quadratic production volume over production intervals. When the survival method was used, the authors found "saucer shaped" average cost curve similar to that reported by Stigler (1958) or Scherer, et al (1975). The cost function approach found that larger scale refineries (those with production up to 500,000 barrels per day) exhibited decreasing production costs.

Hibdon and Mueller used data from 1947-1984. The trends over the last 25 years seem to support Hibdon and Mueller's findings and conclusions. The number of operable crude oil refineries in the United States decreased from 301 to 149 between 1982 and 2003. The distillation capacity of U.S. refineries decreased only slightly from 17.89 million to 16.76 million barrels per day over the same period. In 1982, 129 of the 301 operable refineries were located in Gulf Coast states and Oklahoma. By 2003, only 57 refineries operated in these states. Distillation capacity in these states was 8.7 million barrels per day in 1982 and fell to 8.2 million barrels per day in 2003.

While distillation capacity of the Gulf Coast region decreased by 5.7 percent, the number of production workers decreased by 45 percent and production hours decreased by 37.5 percent. Real wages varied by state. The average hourly wage of a production worker in Louisiana increased from \$14.34 to \$26.62. In Texas, the average real wage increased from \$13.04 to \$27.82. Similar increases occurred in Oklahoma with the average real wage increasing from \$12.60 to \$22.38.

The overall trend for the crude oil refining industry in this region and the U.S. as a whole is to increase the average refinery production and reduce both the number of production workers and production hours. This chapter estimates the demand for labor in the petroleum refining industry for the Gulf Coast region and attempts to explain the mechanism through which labor reduction is taking place. The possible alternatives are that the increasing price of labor relative to capital caused substitution between the inputs, technological progress has reduced the quantity of labor needed to refine a given quantity of crude oil, or that further consolidation has occurred as a result of the decreasing costs within the industry.

The model will be estimated based on the duality conditions of the cost-minimization problem. The first-order condition for the cost-minimization problem provides the conditional derived demand for labor. Estimated coefficients of this demand equation can be used to estimate the underlying production technology.

There are two common methods of estimating factor demand models. The first is to use national level data for several industries. The second is to use a panel of several firms to estimate the factor demands for a specific industry. National data are generally considered to be the less attractive of these alternatives because aggregation errors bias the parameter estimates.

This chapter uses state level and Petroleum Administration for Defense District (PADD) level data to estimate labor demand. These data may be subject to aggregation errors but considerably less error than national data.

It also provides information not available within firm level data. This stems from differences across states which may be aggregated within the firm level data. For example, if a firm operates in Louisiana and Texas and the data provide total worker hours and total capital expenditure for this firm, these data are subject to aggregation bias because they do not account for differences in labor markets, state laws or resource endowments that affect quantity produced, the input mixes used to produce that quantity or the factor prices in each location. It only tells the researcher the total quantity, total input mix, and average factor price across the firm.

Petroleum and chemical manufacturing have some unique characteristics which state level data may be better suited to address. First, production tends to occur in close proximity to the resource. Because labor markets are segmented across states, if wage rates differ, input mixes should differ as well. Second, the refining process in these industries involves emitting several pollutants regulated by the Environmental Protection Agency. EPA guidelines establish acceptable emission levels but leave compliance to the states. Therefore, there is potential for cross state variation in environmental laws. These differences should result in different production technologies and different optimal output levels in each state. These sources of cross state variation are likely to be time invariant and make panel data methods ideal for estimation.

The paper begins with a description of the refining process. Once the process is described, I provide the theoretical foundation for the model. This is followed by empirical results and a conclusion.

The Refining Process

Crude oil consists primarily of hydrogen and carbon. It is not of uniform consistency. The color is anywhere from clear to dark black and the viscosity can be close to that of water or tar like in thickness. The common measure of crude oil density is API gravity. Oil with a gravity of 10 has the same density as water. The density decreases as the API gravity increases. Oils with gravity greater than 10 are less dense than water. Oil is usually categorized as light, medium, or heavy. Light oils have gravities above 31.1 API. Medium oils have gravities between 22.3 and 31.1 API and heavy oil has gravity below 22.3.

Crude oil has no direct uses. It must be refined to produce such commodities as asphalt, heating oils, motor fuel or waxes. Liquid crude oil is pumped into a furnace heated to about 600^{o} C, which transforms the petroleum into various gases state. The gases are then released into a heated distillation tower. The temperatures within the tower decline as elevation increases. As the gases rise through the tower they eventually reach an elevation where the column temperatures are below the boiling point for the particular hydrocarbon fraction. When this occurs the hydrocarbon condenses to liquid and is captured by a tray that removes it from the column.

Once the fraction is removed from the column there are two alternatives that can occur. The first is it is sent to a cooling area to further condense at which point the refining process is complete and the end product is stored. The second alternative is that further chemical processing occurs. The reason for further chemical processing is to convert hydrocarbons with lower levels of demand into products with greater demand particularly to convert other hydrocarbons into motor gasoline.

There are three categories of chemical processing. The first is cracking which breaks longer-chain hydrocarbons into shorter ones. This is done with heat (thermal cracking) or a catalyst (catalytic cracking). Thermal cracking involves heating the large hydrocarbons to high temperatures and often uses high pressure as well. The three types of thermal cracking are steam, visbreaking, and coking. The steam method uses temperatures around 1500^{o} F to break ethane, butane and naphtha into ethylene and benzene which are used in chemical manufacturing. Visbreaking heats the residual substances in the distillation tower to about 900^{o} F in order to burn any gas oil in the residual. This process yields tar. Coking heats the residual to temperatures above 900 F until it cracks into heavy oil, gasoline and naphtha. The remaining residue is coke, an almost pure carbon which is pure carbon.

Catalytic cracking uses catalyst to increase the speed of the cracking process. The catalysts include zeolite, aluminum hydrosilicate, bauxite, and silica-alumina. Fluid catalytic cracking heats a fluid catalyst and the hydrocarbons to about 1000^{o} F to crack heavy gas oils into diesel oils and gasoline. Hydrocracking converts heavy oil into gasoline and kerosene. It uses high pressure, hydrogen gas, and temperatures below 1000^{o} F. Once cracking is complete the hydrocarbons are sent through another distillation tower to separate the new fractions.

The second method of chemical processing is unification. This process combines smaller hydrocarbons into larger ones. The most common method is called catalytic reforming. It uses a catalyst to combine naphtha into aromatics which are either used to make chemicals or as blenders in gasoline. A by-product of this process is hydrogen gas which can be used for hydrocracking.

Alteration rearranges the molecular structure in one fraction to produce another. The most common process is called alkylation. In alkylation compounds with low molecular weights such as butylenes are mixed with a catalyst such as sulfuric acid. The resulting products are high octane hydrocarbons used in gasoline blends to reduce engine knocking.

Once distillation and chemical processing is complete the processed fractions are treated to remove impurities such as sulfur, nitrogen, water, metals and salts. Treatment occurs through three steps. First, the fractions are passed through a column of sulfuric acid to remove unsaturated hydrocarbons, nitrogen compounds, oxygen compounds, and solids such as tars and asphalt. Second, the fractions are sent through an absorption column to remove water. Finally, the fractions are passed through scrubbers to remove sulfur compounds. Once treatment is complete blending takes place to make final products.

The Model

The objective of this research is to estimate the conditional derived demands for factors of production in petroleum refining. First, consider the factors that are most important in refining. Capital and labor are standard factors. The description of the refining process shows that the distillation and chemical processing stages use extreme heats primarily reached through burning gas and recycling steam. No purchased energy source accounts for more than 30% of the total energy. Therefore, I neglect the impact of energy on the production technology and use the standard Cobb-Douglas production function for a representative firm

2.1 $$q_{it} = (k_{it})^{\alpha}(l_{it})^{\beta}\exp(\gamma t)$$

where q measures the output of the refining process, k represents capital, and l labor inputs for plant i at time t. γt represents growth from factor neutral technological advance. This production function is a constraint to the cost-minimization goal of the plant. The resulting objective function is provided by equation 2.2.

2.2 $$\min_{K,L,}\; rk + wl + \lambda(q - k^{\alpha} l^{\beta} \exp(\gamma t))$$

where r, and w are the predetermined input prices of capital and labor.

The first-order conditions yield logarithmic derived demands. They are defined in log form in equations 2.3 -2.4.

2.3 $$\ln l^{*} = \gamma_L + \frac{1}{\alpha+\beta}\ln q + \frac{\alpha}{\alpha+\beta}\ln(r/w) - \frac{\gamma}{\alpha+\beta} time$$

2.4 $$\ln k^{*} = \gamma_K + \frac{1}{\alpha+\beta}\ln q + \frac{\beta}{\alpha+\beta}\ln(w/r) - \frac{\gamma}{\alpha+\beta} time$$

Refined output is obtained from the Energy Information Administration (EIA). They do not actually measure the output produced by the refineries. Instead, they measure the millions of barrels of crude oil and blenders used as inputs. The law of conservation of mass implies that there will be a one to one correlation between barrels of crude oil inputs and barrels of refined output.

The EIA data are available for Petroleum Administration for Defense Districts (PADDs) and subsets thereof. Two of the subsets are Texas inland and Texas coastal. Refined petroleum production for Texas is obtained by summing the outputs for these two subsets. It is not possible to derive outputs specifically for Oklahoma or Louisiana from the EIA data. Oklahoma is in a sub-district that includes Kansas, Missouri, and Nebraska. The ideal solution would be to estimate the derived demands for this region; however, continuous data on wages are unavailable for both Kansas and Missouri (no refining takes place in Nebraska although it is part of the PADD). Instead, the PADD output is used as a measure of output for Oklahoma causing some measurement error.

Louisiana is separated into parts of two PADD sub-districts. The first accounts for production along the Louisiana coast as well as the coastal regions of Alabama and Mississippi. The second accounts for production in Northern Louisiana, Arkansas, and various counties in Alabama, and Mississippi. The data for these two sub-districts is summed to measure the output for Louisiana. The wage and hours data for Arkansas, Mississippi and Alabama are incomplete and measurement error will influence these estimates as well.

Equation 2.4 describes the demand for capital services. It is difficult to measure capital because it is heterogeneous. Furthermore, most capital involved in refining is structures such as distillation towers and catalytic cracking units. Also, no new refineries have been built in the United States since Marathon built a facility in Garyville Louisiana in 1976. These facts make an accurate measure of capital difficult to obtain. Therefore, capital demand is not estimated.

The rental rate of capital is derived from the Producer Price Index for capital goods. The derivation used can be found in Jorgenson (1963). The formula is

2.5 $$r_t = P_t{*}i_t + d_t - (P_{t+1} - P_t).$$

This equation states that the rental rate of a unit of capital is the opportunity cost of forgone investments and the depreciation of the capital minus the change in value due to annual inflation. I assume that ex-ante capital investment is mobile across states, so the capital rental rate is approximately constant across states.

Labor input is measured as production worker hours. Production worker wage rates are calculated by dividing total compensation to production workers by the total number of hours worked. These data are measured at the state levels by the Census of Manufactures for the years 1987, 1992, 1997, and 2002. For all other years labor input is obtained from the Annual Survey of Manufactures.

The model uses the previously defined data for the years 1982-2003 to estimate the derived demand for labor. The general form of the model is derived from equation

2.3. I assume that l^*, the ideal demand for labor for a plant, is not perfectly attained in any period so that actual labor l is defined as $l = l^*\exp(\varepsilon)$. This results in equation 2.6

2.6 $$\ln l_{s,t} = \gamma_L + \frac{1}{\alpha+\beta}\ln q_{s,t} + \frac{\alpha}{\alpha+\beta}\ln(r_t / w_{s,t}) - \frac{\gamma}{\alpha+\beta} time + \varepsilon_{s,t} \quad .$$

The data for this chapter are state level and our demand equation is derived from a cost-minimization problem for a plant. We estimate two models based upon different assumptions of the firms. The first model assumes that each state acts as a firm. Under this assumption, equation 2.6 is estimated as

2.7 $$\ln L_{s,t} = \gamma_L + \frac{1}{\alpha+\beta}\ln Q_{s,t} + \frac{\alpha}{\alpha+\beta}\ln(r_t / w_{s,t}) - \frac{\gamma}{\alpha+\beta} time + \varepsilon_{s,t}$$

where lnL represents total production hours worked in a state and lnQ represents total crude production in a state.

An alternative model uses the assumption that all firms within a state have the same production function and hire resources at the same prices. An implication is that an average plant is representative of all plants within the state. This means that average plant output is representative of all plants and average labor hours per plant is representative of labor hours for any plant. Using this assumption, the derived demand for labor of a representative plant is

2.8 $$\ln l_{s,t} = \gamma_L + \frac{1}{\alpha+\beta}\ln \overline{Q}_{st} + \frac{\alpha}{\alpha+\beta}\ln(r_t / w_{s,t}) - \frac{\gamma}{\alpha+\beta} time + \varepsilon_{s,t}$$

where $\overline{Q}_{st}$ is the average output per plant in state s at time t. The total demand for labor in state s at time t is the product of the number of plants in the state and the demand for labor at a representative plant. The total labor demand equation based on this assumption is described in equation 2.9.

2.9 $$\ln L_{s,t} = \gamma_L + \ln N_{s,t} + \frac{1}{\alpha+\beta}\ln \overline{Q}_{st} + \frac{\alpha}{\alpha+\beta}\ln(r_t / w_{s,t}) - \frac{\gamma}{\alpha+\beta} time + \varepsilon_{s,t}$$

Under this assumption, aggregate conditional demand is determined by the average production of a plant, the relative prices of inputs and the total number of operable plants $N_{s,t}$ in state s on January 1 of year t.

The inverse of the coefficients of $\ln Q$ and $\ln \overline{Q}$ provide estimates of returns to scale under the assumptions of the models. It is important to note that even if these estimated coefficients are biased the returns to scale estimate will be consistent. The estimate should be approximately one if the production function exhibits constant returns to scale and greater than one if the production function exhibits increasing returns to scale. Since petroleum refining involves crude oil being piped throughout the refinery, it is possible that increasing the size of the pipes increases the volume of oil being processed and therefore increasing returns cannot be excluded as an alternative.

Furthermore, the consolidation of crude oil refineries and labor reductions are similar to those of copper refining described by O hUllachain and Mathews (1996). The authors concluded that copper refining was restructured due to cost reductions associated with economies of scale. $\frac{\alpha}{\alpha+\beta}$ is the capital share of production.

Results

First, equations 2.7 and 2.9 are estimated separately for each state. Next, the equations are estimated with the pooled data. Augmented Dickey Fuller tests for each time series provided mixed results. No series was stationary for all states. Some series were stationary for each state but the series varied by state. The Maddala and Wu (1999) test procedure failed to reject the null hypothesis that times series of all cross sections contained a unit root for each series used in the regression equation. Therefore, we conclude each time series is nonstationary.

OLS estimates of models with nonstationary variables may lead to spurious inferences about the relationships between variables. Kao and Chiang (2000) shows that the probability of finding spurious relationships diminish in panel data models as the number of cross sections increases. However, our panel consists of three cross sections so spurious inferences must be a concern.

For the single state equations, we compare results of equations estimated in first differences to equations estimated in levels. For the pooled data, we compare results of three estimation techniques. First, we use least squares dummy variables to estimate the models in levels. Next, we estimate the panel in first differences. Finally, we use a technique proposed in a working paper by Hu (2008). Her procedure applies a Cochrane Orcutt correction for each state in the panel equation. She shows that when cointegration exists, Cochrane Orcutt corrected estimates are at least as efficient as OLS estimates. When there is no cointegrating relationship between variables, the correction method provided consistent estimates with faster convergence than any other method examined

The intuition of this approach is that states with cointegrating relationships among regression variables will have stationary random disturbances. First differencing these variables causes a loss of information and a loss of efficiency. A Cochrane Orcutt transformation of these variables corrects for serial correlation without first differencing the variables. For states where no cointegrating relationships exist between the variables, the Cochrane Orcutt procedure determines that the autocorrelation is one and the correction is equivalent to first differencing. This procedure differences the nonstationary cross sections of the regression equation and estimates the cointegrated cross sections in levels.

Table 2.1 compares the regression estimates of equations 2.7 and 2.9 for Texas. Tables 2.2 and 2.3 provide comparisons of estimation results for Louisiana and Oklahoma, respectively. Regression equations 2.7 and 2.9 provide little information when estimated separately for each state. For Texas, both the levels and first difference estimation of equations 2.7 and 2.9 provide nonsensical results. The estimated coefficients of lnQ and lnQavg are negative and insignificant in levels. An F-test of the first differenced equations fails at the 10 percent level to reject the hypothesis that all coefficients are jointly zero.

Table 2.1. Estimates of Equations 2.7 and 2.9 for Texas

	Levels Estimate	s.e	First Differences Estimate	s.e
lnQ	-0.412	0.590	0.249	0.416
ln(pcap/wage)	0.654	0.207	0.231	0.137
time	0.008	0.012	-0.019	0.014
R^2	0.813		0.229	
lnQavg	-0.632	0.412	0.093	0.403
lnNumber	-0.026	0.418	0.394	0.402
ln(pcap/wage)	0.406	0.154	0.215	0.130
time	0.021	0.009	-0.004	0.015
R^2	0.915		0.12	

For Louisiana, the estimates of lnQ and lnQavg are positive but insignificant in both levels and first differences. The models all provide significant coefficient estimates of capital share that range from 0.24 to 0.27. However, the first differenced estimates of equation 9 failed at the 10 percent level to reject the hypothesis that all coefficient estimates are jointly zero.

Table 2.2. Estimates of Equations 2.7 and 2.9 for Louisiana

	Levels Estimate	s.e	First Differences Estimate	s.e
lnQ	0.310	0.774	0.634	0.452
ln(pcap/wage)	0.273	0.104	0.239	0.090
time	-0.004	0.013	-0.015	0.015
R^2	0.511		0.295	
lnQavg	0.294	0.766	0.665	0.467
lnNumber	0.605	0.805	0.547	0.498
ln(pcap/wage)	0.261	0.104	0.237	0.092
time	0.003	0.014	-0.020	0.018
R^2	0.47		0.286	

Table 2.3. Estimates of Equations 2.7 and 2.9 for Oklahoma

	Levels Estimate	s.e	First Differences Estimate	s.e
lnQ	1.530	0.272	1.024	0.356
ln(pcap/wage)	0.240	0.057	0.175	0.109
time	-0.023	0.003	-0.024	0.016
R^2	0.928		0.339	
lnQavg	1.116	0.361	0.631	0.457
lnNumber	1.321	0.289	0.907	0.360
ln(pcap/wage)	0.257	0.055	0.196	0.108
time	-0.011	0.008	-0.009	0.019
R^2	0.938		0.401	

F-tests of the Oklahoma equations 7 and 9 rejected the hypothesis that coefficient estimates are jointly. The first differenced estimate of lnQavg was 0.63 but insignificant. The differenced estimate of lnQ is 01.02 and significant. The capital shares were estimated as .18 and .20 for equations 7 and 9 respectively. Both estimates are significant at the 10 percent level.

The results of estimating equations 2.7 and 2.9 by pooling the time series for all three states are reported in Table 2.4. The pooled results are more informative than the estimation results reported in Tables 2.1 -2.3. This is consistent with findings by Baltagi, et al (2000) that pooling data provides more reasonable and precise coefficient estimates.

The estimates of equation 2.7 are quite informative. The coefficient estimates of lnQ are all significant and range from 0.76 to 0.9. The point estimates of returns to scale range from 1.11 to 1.32. However, these estimates are too imprecise to make reasonable inferences about the returns to scale. I cannot reject a null hypothesis that returns to scale equal to 1 and I cannot reject a hypotheses that returns to scale equal 1.5 or 2. The estimates of capital share range from .21 to .31 and estimates of technological progress indicate that the demand for labor decreases between 1.7 and 2.3 percent annually, holding output and relative prices constant.

Table 2.4. Panel Estimates of Equations 2.7 and 2.9

	Levels Estimates	s.e	First Differences Estimates	s.e	Cochrane Orcutt Correction estimates	s.e
lnQ	0.904	0.233	0.756	0.214	0.860	0.243
ln(pcap/wage)	0.313	0.053	0.212	0.058	0.254	0.048
time	-0.017	0.003	-0.023	0.008	-0.018	0.003
R^2	0.851		0.258		0.999	
lnQavg	0.342	0.251	0.548	0.232	0.400	0.234
lnNumber	0.781	0.210	0.770	0.209	0.704	0.273
ln(pcap/wage)	0.311	0.048	0.216	0.057	0.282	0.048
time	0.003	0.006	-0.011	0.010	-0.003	0.006
R^2	0.846		0.306		0.999	

The coefficient estimates of equation 9 are also more informative for the full panel. The estimates of lnQavg range from .34 to .57, indicating that estimated returns to scale range from 1.75 to 2.94. These estimates are also too imprecise to make reasonable inferences of returns to scale. Null hypotheses of returns to scale of 1, 1.5, or 2 could not be rejected. The range of capital shares was 0.21 to 0.31. The estimates of technology were not significant in equation 9. The estimates of the labor demand elasticity with respect to the number of plants ranged from .70 to .78. All plant elasticity estimates were significant.

Conclusion

Coefficient estimates from equations 2.7 and 2.9 provide much information about the refining industry. Both equations provide capital share estimates between 0.21 and 0.31. Moroney (1972) found similar estimates for capital, by directly estimating coefficients of a Cobb-Douglas production function for crude oil and coal production. The estimates of equation 2.7 seem to indicate that technological change is reducing the demand for labor between 1.7 and 2.3 percent annually. Equation 2.9 finds that this change in the demand for labor is accounted for by increasing the volume of operations in existing plants and reducing the number of plants. The coefficient estimate for technological improvement is not significant in equation 2.9.

We cannot conclusively determine whether the reduction in labor hours is the result of economies of scale. It is clear, however, that the demand for labor is gradually decreasing in the refinery industry and the size of refineries is increasing. This suggests that much of the reduction in production hours is the result of movements along the isoquant resulting from increases in the relative price of labor over time.

CHAPTER III

A MODEL OF U.S. OIL PRODUCTION, 1981- 2004

Introduction

In February 1999, the national average price for a gallon of unleaded regular gasoline was $0.96. By June 2000, the time of the Democratic and Republican National Conventions, the price jumped to $1.62 per gallon. The timing of the sharp price increase was among several factors renewing interest in U.S. energy policy. In the resulting debates, both sides stressed the economic risks associated with energy price volatility and believed that the U.S. needed to reduce dependence on oil.

The central question is how to achieve this reduced dependence. U.S. consumption is approximately 20.5 million barrels per day while production is about 7.5 million. Major oil and drilling companies such as British Petroleum, and ExxonMobil, among others, suggest increased exploration and further development of domestic reserves. They argue that subsidies, tax credits, and development on federal lands will significantly reduce U.S. dependence on foreign oil.

Environmental groups such as The Natural Resource Defense Council and The Sierra Club believe instead the U.S. should stress policies that reduce the demand for petroleum products. They argue that supply-based initiatives will be ineffective because there are insufficient reserves left in the U.S. to meet the current demands. Domestic demand is growing, so domestic supply would have to increase even faster in order to reduce imports. Drilling, shipping and consumption of petroleum products also have negative environmental impacts, ranging from increased greenhouse gas emissions to degradation of pristine wildlife habitats. The benefits of increasing supply, in their opinions, are negligible in comparison to the costs.

This paper attempts to model U.S. supply. In doing so, we assess claims that increasing supply is a feasible way to reduce U.S. dependence on foreign oil.

To our knowledge, there are two recent publications looking at the impacts of reserve or supply increases. The first is an EIA (2000) assessment of production from the Arctic National Wildlife Refuge (ANWR). Within ANWR is a potentially large oil reservoir along the coastal plain. In section 1002 of the Alaska National Interest Lands Conservation Act of 1980, Congress deferred land management decisions regarding oil and natural gas production within this area. The USGS collected and surveyed existing geological data to estimate the potential oil available from the 1002 region of ANWR. The USGS reports estimates of the 95%, 50%, and 5% likely reserve quantities. The EIA used these USGS estimates as the basis for estimating recoverable reserves. Based upon estimated recoverable reserves, EIA simulated development and production within 1002, using the 50% likely reserves of 10.3 billion barrels. This simulation projects peak production of 876,000 barrels daily in 2024. This report is specific to the ANWR region and does not generalize to the entire U.S.

Kaufman and Cleveland (2001) performed a second study by applying a Vector Autoregressive model of prices, proved reserves, production, and rationing by the Texas Railroad Commission to aggregate U.S. data. The authors conclude that development of ANWR would do little to curb U.S. dependence on imported oil.

Our model, unlike those previously mentioned, uses panel data from 23 states and the Federal Gulf of Mexico. These data include interstate differences in prices and reserves that are unavailable with aggregate data. Panel data can also account for time-invariant factors such as geology or geography that may influence production.

Section I discusses the theoretical model that serves as the basis for our econometric estimates. Section II describes the data and their time series properties. The econometric models are specified in section III and section IV discusses major results. Section V applies these results for increasing domestic oil production. Section VI summarizes and concludes the paper.

Theory

Pindyck (1978) presented the theoretical model used in this study. It is a valuable generalization of Hotelling's (1931) model, which has the restrictive assumption that at each moment in time the stock of an exhaustible resource is fixed and known with certainty. In this setup, the goal of competitive producers is to extract the resource to maximize the present value of their firm's profits. Pindyck's substantive contribution is to allow firms to devote effort both to production and to exploration for new reserves, changes to cumulative reserves, $\dot{x}$, depend upon the effort, w, to find new sources of oil and the quantity of cumulative reserves, x, already discovered (i.e. $\dot{x} = f(w, x)$). It is assumed $f_w > 0$ (greater exploratory effort increases reserves) and $f_x < 0$ (the larger the volume of cumulative reserves discovered, the smaller the volume of reserves added due to reserve depletion). The firm now chooses the amount of effort and the quantity of oil production that maximize the present value of profits over the life of the firm.

Equation (1) formally expresses the model.

3.1 $$\underset{q,w}{Max}\ \mathrm{W}= \int_0^{\infty} (p(q)q - c_1(R)q - c_2(w))e^{-\delta^* t}dt$$

subject to

3.2 $$\dot{R} = \dot{x} - q\,,$$

3.3 $$\dot{x} = f(w, x)\,,$$

and

3.4 q, R, w, x being nonnegative.

where p(q) is a price function, q is the firm's production, $c_1(R)q$ is the production cost which depends on reserves R and the level of production and $c_2(w)$ is the portion of costs that depend on exploratory and development efforts and δ is the risk-free rate of return.

Equation (1) allows firms to be either perfect or imperfect competitors (that is $\frac{\partial p}{\partial q}$ may be negative or zero). We assume prices are determined in a world market and all U. S. firms are competitive price takers. This assumption is consistent with empirical findings by Griffin (1985). We assume $c'_1(R) < 0$. That is, as reserves increase production costs decrease. This assumption is consistent with facts concerning lifting costs. Other things equal, the more oil within a reservoir the greater the reservoir pressure. The pressure acts as a "natural lift" and reduces the pumping necessary to extract the oil. As the reservoir is depleted (i.e. R decreases) artificial lift (and greater production cost) is needed to extract the same quantity of oil.

In addition to an effort cost $c_2(w)$ now being part of the model, the constraints facing the firm now change as a result of removing Hotelling's fixed stock assumption. These are reflected in 3.2 and 3.3. With a fixed stock, S, the constraints would simply be $\dot{S} = -q$, implying that any production reduces the stock by the quantity produced. By relaxing this assumption, reserves increase because of exploratory effort and decrease because of production. As stated previously, cumulative reserve additions depend upon w and x.

We can incorporate 3.1 – 3.4 as a Hamiltonian to examine the determinants of optimal production and effort. The Hamiltonian is equation 3.5.

3.5 $$H = (pq - c_1(R)q - c_2(w))e^{-\delta^* t} + \lambda_1(f(w,x) - q) + \lambda_2(f(w,x))$$

Two of the four first-order conditions for maximization of 3.5 relate to production. The first-order conditions are

3.6 $$H_q = (p - c_1(R))e^{-\delta^* t} - \lambda_1 \geq 0$$

and

3.7 $$H_R = -c_1'(R)qe^{-\delta^* t} = -\dot{\lambda}_1.$$

There are two first order conditions we exclude. The excluded equations involve the optimal level of exploratory effort. Exploratory effort affects optimal production by determining the volume of proved reserves available for production.

Equation 3.6 shows the necessary conditions for optimal production. The determinants are the market price, reserves, and the shadow price, λ_1, of forgoing production until a future period. Rewriting 3.6 as $(p - c_1(R))e^{-\delta^* t} \geq \lambda_1$ indicates that firms should produce as long as the present value of profit on the marginal unit is greater than or equal to the benefit of withholding that unit for future production. Optimal production occurs when 3.6 holds with equality. If 3.6 holds with inequality, the firm faces a capacity constraint that limits maximum output.

Equation 3.7 tells us the shadow price of holding the marginal unit decreases over time. This occurs because lifting costs increasing as reserves decrease ($c_1^{'}(R) < 0$). Therefore, production will generally cease before reserves are exhausted.

By combining equations 3.6 and 3.7 with the first-order conditions of the Hamiltonian with respect to effort and cumulative reserves, this model shows that low initial reserves and high initial effort result in cumulative reserve and production paths that are consistent with the "peak" models used by Hubbert (1962, 1967, 1982). Therefore, this model provides an economically based explanation for the production path actually observed within the United States. Nonetheless, econometric models need not be restricted to the symmetric production paths embedded in Hubbert type models.

Oil production models based upon prices and reserves as well as political constraints tend to explain current production and forecast future production better than purely economic or purely geological models. (See Moroney and Berg (1999), Cleveland (1991), Kaufman (1991), Cleveland and Kaufman (1991, 2001))

In general terms, the models used in this study are based upon 3.6; where, $q = f(R, p) + \varepsilon$. Section 3 provides the specific models. The data are described next.

Data

The data are state-level time series for oil production, proved reserves, and domestic first purchase prices. They are available on the Energy Information Administration website. Each time series spans 1981-2004. These data series are available for 23 states. Additionally, data are available for the Federal Gulf of Mexico region from 1986-2004. Proved reserves and production are measured in thousands of barrels annually. Domestic first purchase prices are converted from current year dollars per barrel to 1982 dollars per barrel using the Producer Price Index for final goods. The Energy Information Administration does not report the 1990 price for West Virginia. Instead, the 1990 average price for the Petroleum Administration for Defense District I, which includes West Virginia, is used. Over the sample period, there is usually no more than a \$0.25 difference in the PADD I price and the West Virginia price.

Production is the volume of crude pumped annually. Proved reserves must satisfy U.S. Securities and Exchange Commission guidelines as estimates of the quantity of oil that can be produced in light of current technology and economic conditions.

The domestic first purchase prices are prices for an arms-length transaction at the wellhead. The producer is indifferent between shipping the oil to market and selling it at the wellhead. Accordingly, these prices account for interregional heterogeneity in oil quality and transportation costs.

These data cover an overwhelming majority of all production within the United States. In 2004, U.S. oil production was 1.98 billion barrels of which these data account for 1.947 billion. The only area of note that is excluded is the Federal Offshore Pacific

Region, which accounted for about 1% of U. S. production and 2.5% of U. S. Reserves in 2004.

Unlike a regression estimated using time series data for one cross-section, estimates from models with panel data will not show spurious correlations even if the time series for the variables are non-stationary. However, if non-stationary variables are cointegrated, Least Squares Dummy Variable coefficient estimates will have a non-negligible bias. The unit root tests in panel data are similar to those for a single time series. The most commonly used procedures are those proposed by Levin and Lin (1999), Im, Pesaran and Shin (1997) and Maddala and Wu (1999).

These tests are based on an Augmented Dickey Fuller specification such as 3.8

3.8 $$y_{j,t} = \alpha_j + \rho_j y_{j,t-1} + \delta_j t + \sum_{i=1}^{L} \gamma_{ij} \Delta y_{j,t-i} + \varepsilon_{j,t}$$

All three tests have a null hypothesis $\rho_j = 1$ for all j. The alternative hypothesis differs for each test, as do the test statistics.

Levin and Lin pool the cross-sections to estimate 3.8 via OLS where $\rho_j = \rho$ for all j cross-sections. The authors show that the asymptotic distribution of ρ has zero mean but the variance depends on the specification of trends, and fixed effects. This test's alternative hypothesis is that $\rho < 1$. The test is restrictive because it assumes ρ is the same for all cross-sections. Failure to reject the null hypothesis means either there is a unit root or the stationary autoregressive process is not the same for all cross-sections.

Im, Pesaran, and Shin (1997) estimate 3.8 or some variation that excludes constants and/or time trends separately for each cross-section and then averages the t-statistics for each cross-sectional unit root test. The authors use sequential limit theory to show that if the t-statistic for each cross-section is identically and independently distributed, under the null of a unit root, the average of the t-statistics converges in probability to a distribution based on the underlying Weiner processes. As N tends to infinity, the test statistic converges in distribution to standard normal. The alternative hypothesis is that $\rho_j < 1$ for at least one cross-section. The test is not as restrictive as the Levin Lin procedure, but rejecting the null does not imply all cross-sections of the tested variable are stationary. Also, when the null is rejected, the test does not identify the stationary and non-stationary cross-sections.

Maddala and Wu (1999) proposed the third procedure. The procedure uses 3.8 or some variation similar to Im, Pesaran, and Shin. Instead of using the t-statistics, this test uses the p-values associated with the t-statistic. The p-value of the null $\rho_j = 1$ is denoted p_j for each cross-section j. Maddala and Wu show that the test statistic $-21n(p_j) \sim X^2(2)$. Therefore, the test statistic, $-2\sum_{j=1}^{N}\ln(p_j) \sim X^2(2N)$, under the null.

The advantage of this procedure is that any standard unit root test can be performed. The lag, trend structure and number of observations can vary by cross section. The disadvantages of this test are that a rejection of the null indicates that the series is stationary for at least one cross-section but does not identify which cross-section(s) are stationary. Additionally, Maddala and Wu report that the power of this test diminishes when there is contemporaneous correlation among cross-section observations.

After proposing the p-value based test, Maddala and Wu (1999) performed Monte-Carlo simulations to compare the accuracy in terms of size and power of their test to the Levin and Lin and Im, Pesaran and Shin procedures. For a sample of 25 cross-sections with 25 time periods, (the sample size most similar to the sample used within this study), the authors found their test was the most accurate in terms of Type I errors but that none of the tests had power above 10%. The implication is that for samples of this size these tests almost never reject the null of $\rho = 1$ even if the true $\rho = 0.9$. Furthermore, the power of all these tests diminishes if the cross-sections are contemporaneously correlated.

Tables 3.1a – 3.1c show the results of these tests for the lnProduction, lnPrice, and lnReserves series in levels. The tests were performed using E-Views built in programs with between 1 and 4 determined by the Schwartz Information Criteria. All tests of first differenced variables rejected the null of a unit root.

These tests provide mixed results when the series are measured in levels but are interpreted to indicate that each series is stationary. LnProduction appears to be stationary when tested with no individual effects or time trends. Both the Levin and Lin and Maddala and Wu procedures soundly reject the null of a unit root. (E-Views will not perform the IPS test unless a constant term is included). However, if the lnproduction series is tested with constants terms or constant terms and time trends the tests almost always fail to reject a unit root.

The Levin and Lin procedure rejects the null of a unit root in all specifications of individual effects and time trends for the lnprice series. Maddala and Wu and IPS tests reject the null of a unit root when constant terms are included but fail to reject the null when no constant terms or both constants and time trends are included.

For the lnreserves series, the Levin and Lin test soundly rejects the null of a unit root when no individual effects or time trend are included or when both are included. The specification with individual effects and no time trend is significant at the 0.10 level but barely misses the 0.05 rejection criteria. Maddala and Wu results are similar when no trend and constant are included. Both IPS and Maddala and Wu are similar to the Levin and Lin results when both a trend and constant are included but both tests soundly fail to reject a unit root when only a constant term is included.

Given the low power of these tests and the fact that at least two tests rejected the null of a unit root for some specification of equation 3.8 for each variable, it is reasonable to assume these variables are stationary. If the series are actually non-stationary, Kao (1999) shows that coefficient estimates will converge to zero as the number of cross-sections increases. However, the standards errors converge faster and inference will be incorrect with probability 1 as N tends to infinity. Another property of spurious regression in panel data is that the R^2 of LSDV estimation converges to zero rather quickly. In Monte Carlo simulations of samples with N and T comparable to this study, the average R^2 was less than 0.10. Within this study Kao, proposes four tests for cointegration based upon the residuals of the LSDV estimation. A rejection of the null of a unit root within the residuals indicates that the residual series is stationary. This means either there is a cointegrating relationship between non-stationary variables or that the random error term is stationary because all the variables in the regression are stationary.

Rejection of a unit root is not a sufficient reason to proceed with LSDV estimation. Kao and Chiang (2000) show that the LSDV estimates have a non-negligible bias when cointegrated variables are included. The two standard estimation methods that correct for this bias are Dynamic OLS (DOLS) and Fully Modified OLS (FM-OLS). DOLS estimation, proposed by Kao and Chiang (2000), includes lags and leads of first differenced right hand side variables for each cross-section in order to correct for any endogeneity and/or serial correlation within the random disturbance in the regression.

FM-OLS, proposed by Pedroni (1997), is a two-stage process where the first stage involves LSDV estimation to obtain residuals. The second stage uses the residuals to estimate the long run covariance matrix. The covariance estimation involves a nonparametric kernel density estimator. The estimated covariance matrix is used to perform feasible generalized least squares estimation on the original model. Kao and Chiang (2000) show that DOLS and FM-OLS have the same asymptotic properties but that DOLS provides more accurate coefficient estimates in small samples.

Econometric Model

The unit root tests for ln(production) , ln(price), and ln(reserves) provide mixed results. I proceed on the assumption the series are stationary. If this assumption is incorrect, either (1) the coefficient estimates and R^2 will tend to zero in the case of non-stationary non-cointegrated regression or (2) coefficient estimates will be biased if the variables are cointegrated.

As a further test of the stationarity of the variables, I specify a regression equation as

$$production_{it} = \mu_i + \beta_p price_{it} + \beta_R reserves_{it} + d_1 1986 + d_2 Gulf1 + d_3 Gulf2 + u_{it} \tag{3.9a}$$

Production, price and reserves are measured in natural logarithms. Gulf1, Gulf2 and 1986 are dichotomous variables representing potential shocks to world oil markets that may have influenced domestic production decisions. Gulf1 takes a value of 1 in 1990 and 1991 and 0 in all other years. It represents possible shocks due to the Iraqi invasion of Kuwait and the liberation that ensued. Gulf2 has a value of 1 in 2003 and 2004 and value 0 in all other years. It accounts for the effects of the current gulf war on domestic production decisions.

A dummy variable is included for 1986, a year in which oil prices collapsed because Saudi Arabia changed its role within OPEC. Prior to 1986, Saudi Arabia played the role of "swing producer" within OPEC, and supplied the amount of oil necessary to keep the world price at a target level. However, other OPEC nations exceeded their production quotas, causing Saudi Arabia to produce less oil than it desired. In 1986, Saudi Arabia began a tit-for-tat punishment strategy to enforce the cartel agreement. Under this new strategy, all countries including Saudi Arabia had a production quota. If a country exceeded its quota by X barrels, Saudi Arabia responded by exceeding its quota by X barrels as well. The new strategy flooded the world market, causing the price to fall substantially during 1986.

The residuals of equation 3.9a are used to perform Kao's cointegration tests. These tests reject the null of a unit root; indicating that either the variables are stationary or non-stationary but cointegrated. If the variables in equation 3.9a are cointegrated, the coefficient estimates are biased and DOLS and/or FM-OLS should be performed. I perform DOLS because of its computational simplicity and accuracy in small samples.

The alternative specification of 3.9a under DOLS is

$$\begin{aligned} &production_{it} = \mu_i + \beta_p price_{it} + \beta_R reserves_{it} + d_1 1986 \\ &+ d_2 Gulf1 + d_3 Gulf2 + \sum_{l=-p}^{p} \gamma_{pil} \Delta price_{it+l} + \sum_{l=-q}^{q} \gamma_{Ril} \Delta reserves_{it+l} + u_{it} \end{aligned} \quad (3.9b)$$

This specification includes one lag and one lead of the first differences of prices and reserves (in natural logarithms) for each cross-section. The first-differenced lag and lead variables result in the loss of observations in 1981, 1982, and 2004. For comparison purposes, equation 3.9a, the LSDV estimator, is regressed over the same time-period as 3.9b. Table B3.2 shows the results of these alternative specifications.

The coefficient estimates are very similar across the two equations. Estimated reserve elasticities are 0.80 under DOLS and 0.74 under LSDV estimation. The price elasticity estimates are 0.15 and 0.22 under DOLS and LSDV, respectively. The sign and magnitudes of the dummy variables are also similar. The comparable results between DOLS and LSDV procedures seem to indicate that the LSDV estimates are not biased, as they would be if non-stationary variables are cointegrated. This is taken as further evidence that the production, price, and reserves series are stationary.

Pindyck's theory specifies the factors that determine production but does not provide a functional form. The key variables are prices and reserves. However, the production path also depends on initial reserves and initial effort, data that are unavailable. In order to remedy the influence of initial conditions, I specify the model as

$$\begin{aligned} production^{*}_{it} &= \mu_i + \beta_p price_{it} + \beta_R reserves_{it} + d_1 1986 \\ &+ d_2 Gulf1 + d_3 Gulf2 + \Theta_{1i} time + \Theta_{2i} time^2 \end{aligned} \tag{3.10}$$

Under model 3.10, the ideal level production, measured in natural logarithms, for cross-section i at time t depends on the natural logarithms of prices and reserves in cross-section i at time t, production shocks that are described previously and quadratic time effects that capture the production path implicit by previous investments in effort and development. This model assumes that price and reserve elasticities do not vary across states. However, the μ_i terms allow for state-specific variation in time-invariant factors such as the quality of the oil, or geological structures that can affect extraction costs. The quadratic time variable allows for production differences across states that result from the differences in development and maturity of the states' oil fields. A quadratic time variable permits nonlinear production paths over time. In the Gulf of Mexico, production peaked in 1991 and has declined ever since. In all other states, production has declined at either a constant rate or an increasing rate.

Although 3.10 describes ideal production, firms are unlikely to fully adjust production if they incorrectly forecast prices, revise their reserve estimates, or face an unanticipated shock. Therefore, I use a partial adjustment model to estimate actual production. Under the partial adjustment model, the change in actual production is a fraction of the change of ideal production at time t. That is,

3.11 $$(1-\lambda)(production_{it}^{*} - production_{i,t-1}) + u_{it} = production_{it} - production_{i,t-1}.$$

1-λ is the fraction of ideal production that is achieved within one year, $0 \le \lambda \le 1$. If λ = 1, there is no adjustment towards the ideal level. If λ = 0, full adjustment occurs within one year. Substituting equation (10) into (11) and solving for production yields regression equation (12).

3.12 $$\begin{aligned} production_{it} &= \mu_i^{+} + \beta_p^{+} price_{it} + \beta_R^{+} reserves_{it} + d_1^{+} 1986 + d_2^{+} Gulf 1 \\ &+ d_3^{+} Gulf 2 + \Theta_{1i}^{+} time + \Theta_{2i}^{+} time^2 + \lambda production_{i,t-1} + u_{it} \end{aligned}$$

A "+" superscript over a coefficient indicates a short run coefficient. The long run coefficient is obtained by dividing the short run coefficient by (1-λ). That is, β_p^{+} is the short run price elasticity of production and the long run price elasticity $\beta_p = \dfrac{\beta_p^{+}}{1-\lambda}$.

The LSDV estimation of 3.12 provides estimates that are biased and inconsistent because the lagged dependent variable is used as a regressor. (see Appendix 3.1 for a detailed explanation of the bias). Judson and Owen (1999) examined the bias of the LSDV estimator and the estimation techniques proposed by Anderson and Hsiao (1981), Arellano and Bond (1991), and Kiviet (1995) in samples of comparable size to our data. The authors found that the bias of the LSDV estimator was as much as 20 percent of the true value of a coefficient. Of the alternative procedures reviewed, the method proposed by Kiviet (1995) performed best for samples of this size. (The Anderson and Hsiao and Arellano and Bond procedures are explained in Appendix 3.1).

Nickell (1981) derived the bias of the LSDV estimator. Kiviet's procedure draws upon Nickell's work to derive a consistent estimator of the bias, which is subtracted from the LSDV estimator. Kiviet's method is a two-stage procedure. The first stage involves obtaining estimates using methods proposed by Arellano and Bond or Anderson and Hsiao. The second stage uses these estimates as initial values for an iterative process that minimizes the bias according to Nickell's derivation. Bruno (2005a) has extended Kiviet's procedure to unbalanced panels and written a Stata command (2005b) that can estimate this procedure for both balanced and unbalanced panels.

In addition to the LSDV and biased corrected LSDV estimators, I also estimate equation (3.12) using Arellano and Bond's (1991) GMM method and Anderson and Hsiao's (1981) first-differenced instrumental variables method. These methods are commonly used in the literature but Judson and Owen (1999) find that the Arellano and Bond procedure does little to correct the magnitude of the bias in samples with only a small number of cross-sections. In fact, the LSDV estimator is frequently less biased than the Arellano and Bond estimator when the panel contains a small number of cross-sections. Judson and Owen (1991) find that the Anderson and Hsiao (1981) estimator performs slightly better than Arellano and Bond in panels with few cross-sections but the bias is still non-negligible.

Results

Table B3.3 reports the results for the LSDV, bias corrected LSDV (LSDVc), Arellano and Bond (AB), and Anderson and Hsiao (AH) estimation procedures. The LSDV and Arellano-Bond estimates were very similar. λ is estimated as 0.57 and 0.59 respectively. The short run price elasticity estimates are .06 and .05 and short run reserve elasticity estimates are .08 and .07 respectively. These estimates of λ differ greatly from the Anderson and Hsiao and LSDVc estimates. The AH estimate of λ is 0.89 but not significantly different from 1, implying that no adjustment towards long run equilibrium occurs within one year. The AH short run reserve elasticity estimate of .014 is not statistically significant either. Only the short run price elasticity estimate of .06 is significant.

The LSDVc estimate of λ is 0.77 and differs from the estimates of the other three methods. It is about 25% larger than the LSDV and AB estimates and somewhat similar to the AH estimate. Unlike the AH estimate, the LSDVc estimate is statistically different from both 0 and 1. The LSDVc short run price and reserve elasticity estimates are .06 and .05 respectively. Both are statistically significant and similar to AB and LSDV estimates. The long run price and reserve elasticity estimates are 0.26 and 0.23 respectively. I performed a Wald test to determine that both long run estimates are statistically significant at the .05 level.

The implication of these estimates is that only about 23% of the adjustment towards long run equilibrium occurs within one year and that production has inelastic responses to price and reserves in both the short run and long run. Therefore, any public policies that promote the growth of domestic production through subsidies, reduced royalties, or other means of increasing the price received by producers will have little effect. Policies designed to increase reserves by promoting exploration and development will also have little effect.

Model Simulation 2005 – 2030

To illustrate the point that price and reserve increases will have little effect, I use the LSDVc estimates to simulate production from 2005 - 2030 under three cases. In case 1, the change in proved reserves follows the formula: $Reserves_t = Reserves_{t-1} +$ 1.894 billion – $production_t$. This formula states that current reserves are last years reserves plus the difference between production and gross reserve additions of 1.894 billion barrels. 1.894 billion barrels is the annual average quantity of gross reserve additions in the United States from 1977 to 2005. Price is held constant at $40 per barrel (in 1982 dollars).

In case 2, Reserves follow the same path but an additional 2 billion barrels are added in 2012, the year the EIA estimates that ANWR 1002 production would become available if development starts immediately. In 2013 and 2014, 3 billion more barrels are added in each year and 2 billion barrels of additional reserves are added in 2015. The overall effect is that 10 billion barrels of oil, the USGS mean estimate of recoverable reserves from ANWR, are added over this four year period. In case 2, prices are constant at $40 per barrel as well.

Case 3 allows real prices to grow at 1.5% annually from a base price of $40 in 2005 and allows reserves to change in the same method as case 2.

Figure 3.1 shows the results of these simulations as well as EIA estimates of domestic consumption from 2005-2030. The information provided from these simulations is clear. First, the level of dependence on foreign sources of oil diminishes through the development of ANWR 1002. Second, adding 10 billion barrels to other domestic reserves by 2015 (a nearly 50% increase from current reserve levels) does nothing to reduce dependence on imported oil. Even with this 50% reserve increase attributable to ANWR, the simulation projects domestic production in 2030 to be slightly lower than it is in 2005 (1.83 billion barrels under case 3 vs. 1.89 billion barrels in 2005), while domestic consumption is predicted to increase from 5.55 billion barrels in 2005 to 6.75 billion barrels in 2030. The effect is that net imports increase from 3.66 billion barrels in 2005 to 4.98 billion barrels in 2030 under case 3 or to 5.15 billion barrels under case 1.

Theses simulations suggest that developing and producing reserves from ANWR could postpone the eventual decrease in domestic oil production. But the hypothetical addition of 10 billion barrels from ANWR by 2015, by itself, would not prevent an ever-widening difference between projected domestic consumption and production.

Conclusion

In this study, we have used state-level panel data to estimate US oil production. We find that domestic production responds inelastically to prices and reserves in both the short run and in long run equilibrium. The implication is that policies designed to encourage exploration and development such as tax credits or actions designed to increase net prices received by producers such as subsidies or royalty reductions cannot by themselves to increase domestic production.

Simulations show that adding 1.894 billion barrels to reserves every year and an additional 10 billion barrels over the 2012- 2015 period along with a 1.5% increase in the real price of oil cannot prevent growing dependence on imports until 2030. There are no feasible price increases that can reduce the dependence on foreign oil if consumption grows as the EIA projects.

If energy independence is truly a goal of policy makers, supply-side incentives will have a negligible impact on reaching this goal. These policies may allow production to remain stable near the 2005 levels until 2030 but will not provide the growth necessary to account for the projected increases in consumption.

The reason is reserve depletion. The United States has produced oil for more than a century. Its proved conventional reserves, onshore and offshore combined, reached an all-time peak of 39 million barrels in 1971, and have since dwindled to roughly 21 billion barrels in 2006. Financial incentives assuredly stimulate new exploration and development. But further drilling in well-established areas cannot stem continuing decline in conventional reserves.

CHAPTER IV

U.S. ENERGY DEMAND AND ANTHROPOGENIC CONTRIBUTIONS TO CARBON DIOXIDE EMISSIONS, 1949-2005

Introduction

Among the highlights of President Bush's 2006 State of the Union Address was his declaration that America is "addicted to oil" and the proposed Advanced Energy Initiative that increases Department of Energy clean energy funding by 22% to accelerate research into how Americans power their homes, businesses, and automobiles.

This is but one example of a renewed focus on U.S. energy policy that has emerged since gasoline prices began to rise in the spring and summer of 2000. Others include a proposal by Presidential candidate Al Gore to release part of the strategic petroleum reserve in order to lower the market price of gasoline and other petroleum products and the passage of the Energy Policy Act of 2005 that provides grants and loan guarantees for research into cleaner methods to burn coal, and to further research into biofuels and other technologies that avoid or reduce greenhouse gases.

The two main themes of nearly all energy proposals are national security concerns and environmental concerns. From a national security perspective, volatile energy prices can affect the macro economy (Hamilton 1983, Hamilton and Herrara 2004). There is a potential for supply interruptions due to political hostilities towards the United States from Iran or Venezuela which are both members of OPEC. There is also concern that oil revenues received by other predominantly Muslim nations fund terrorist activities against the U.S. and Western Europe.

Environmental concerns range from the poor air quality associated with sulfur dioxide emissions to the impact of greenhouse gas emissions on global warming. The concerns over global warming were highlighted by the release of the Intergovernmental Panel on Climate Change's Fourth Assessment Report (2007), which concluded that anthropogenic sources of greenhouse gases are "very likely" causes of global warming. The report also concludes that carbon dioxide emissions from burning fossil fuels and changes in land use patterns have the greatest contribution to greenhouse gas emissions. (CO_2 emissions from burning fossil fuels are about 4 times greater than those from changes in land use patterns.)

Given the concerns about energy usage and the environment, we examine economic factors that determine the demand for fossil fuels and the contributions that each fuel makes to human carbon dioxide emissions. We find little substitutability among coal, oil, and natural gas. In fact, it is probably more appropriate to refer to three separate markets than it is to refer to a U.S. energy market.

Previous multi-country panel data studies have shown per capita emissions elasticities with respect to per capita real GDP that are negative or less than one. Holtz-Eakin and Selden (1995) estimated per capita GDP elasticities using quadratic specifications of GDP in levels and natural logarithms. The natural logarithms model estimated the linear coefficient as 0.52. The quadratic coefficient estimate was -0.028. Schmalansee et al (1998) estimated per capita GDP elasticities with a spline regression.

The estimated elasticity was -0.30 for incomes between $9799 and $19627. We use a partial adjustment framework to model both short run and long run effects.

Adjustments to long run equilibrium vary greatly by fossil fuel. Coal achieves full adjustment within one year and natural gas demand adjusts by 50% in one year. The short and long run price elasticities are highly inelastic for each fuel but short and long run real GDP elasticities vary greatly with coal having short run and long run elasticities near 1.40, while the short run and long run elasticities for natural gas are about 0.73 and 1.45, respectively.

The ability to distinguish between the short run and the long run is of paramount importance. Figure 4.1 shows the time pattern on energy consumption per dollar of GDP. There is clearly a downward trend. This has caused some to conclude that increases in GDP lead to a reduction in energy consumption. This conclusion leads to what we believe is an erroneous conclusion that as developing countries continue to grow, they will eventually curb energy consumption and the CO_2 emissions that are a by-product of this consumption. Figure 4.2 shows trends in real GDP and fossil energy consumption over time. There is clearly growth in both time series over the entire sample.

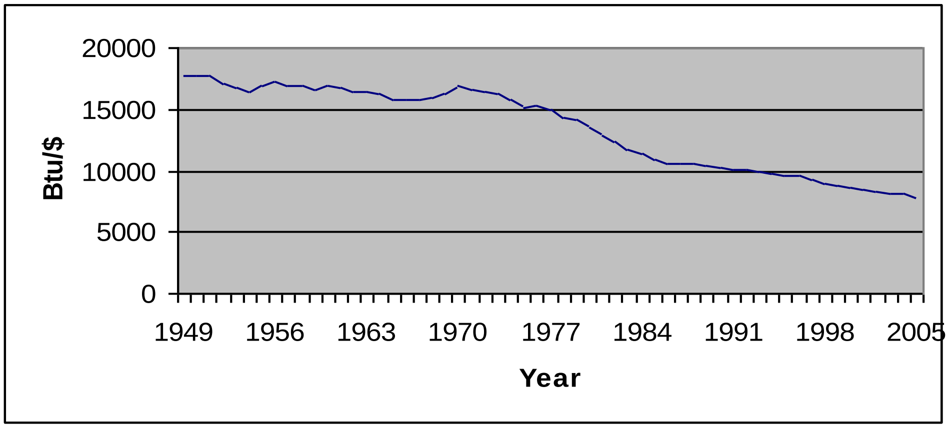

Figure 4.1. Fossil Fuel Consumption/GDP in the U.S., 1949-2005

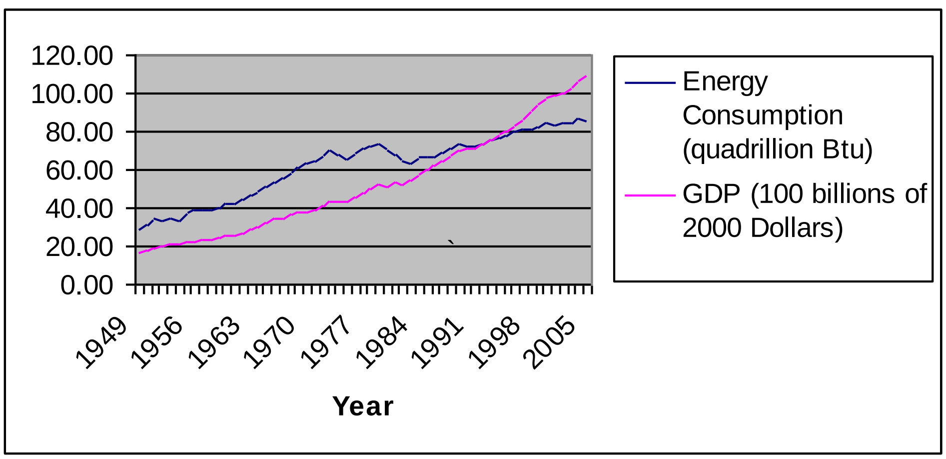

Figure 4.2. Fossil Energy Consumption and GDP in the U.S., 1949-2005

Energy consumption is in quadrillions of Btu. It includes consumption of coal, natural gas, and petroleum. In order to comparably scale GDP, it is measured in 100s of billions. GDP in 1949 is about $1.9 trillion. GDP in 2005 is $11 trillion. GDP is measured in 2000 chain price index deflated dollars.

It is more accurate to say the rate of growth in CO_2 emissions has been less than the growth rate in GDP. We attribute this decline to the differing responses of each fossil fuel to changes in real GDP and the sectors in which growth has taken place.

In 1949, real GDP was $1.64 trillion (in 2000 chain price indexed dollars). There were 29.002 quadrillion Btu of fossil energy consumed. Coal consumption of 11.981 quadrillion Btu was the largest, petroleum was second at 11.883 quadrillion Btu, and natural gas consumption was a mere 5.145 quadrillion Btu. By 2006, real GDP increased almost seven-fold to $11.3 trillion, and fossil energy consumption nearly tripled to 84.76 quadrillion Btu. The energy mix changed substantially. Coal consumption doubled to 22.511 quadrillion Btu while petroleum and natural gas consumption nearly quadrupled to 39.758 and 22.431 quadrillion Btu, respectively.

The growth of the various sectors of the economy is important because each fuel has substantially different CO_2 oxidation rates per Btu and there is virtually no substitution between fuels. The five main sectors are residential, commercial, electricity production, industrial, and transportation.

In 1949, the residential sector consumed about 5.6 quadrillion Btu. About 3.4 quadrillion Btu or 61% of the total was from fossil fuels. The mix was almost evenly distributed with coal, natural gas and petroleum shares of 23%, 18%, and 20% respectively. By 2005, the fossil fuel share of total residential consumption was only 28.2%. Between 1949 and 2005, direct consumption of coal nearly vanished in the residential sector. Gas consumption increased from about 1 quadrillion Btu to 4.5. Petroleum consumption remained about the same.

The pattern is similar in the commercial sector. In 1949, 71 percent of the 3.66 quadrillion Btu of energy consumption came from fossil fuels. By 2005, only 21.5% of the 18 quadrillion Btu of consumption came directly from fossil fuel. Coal consumption decreased from 1.5 to 0.084 quadrillion Btu. Natural gas consumption increased 0.36 to 3 quadrillion Btu and petroleum consumption remained stable.

Although the residential and commercial sectors appear to substantially reduce their usage of fossil fuels, the substitution was towards electricity not towards renewable sources of energy. The electricity sector consumed 4.3 quadrillion Btu in 1949. 46% was from coal, 13% from natural gas, and 9% from petroleum. In 2005, total electricity consumption was 39.7 quadrillion Btu. 51% was from coal, 16% from natural gas and 1.6% from petroleum. Overall, coal consumption increased from 2 quadrillion Btu to 20.5. Natural gas consumption increased from 0.6 to 6.4 quadrillion Btu and petroleum consumption increased from 0.4 to 0.6 quadrillion Btu.

Within the commercial sector, electricity grew from 1 quadrillion Btu in 1949 to 14 quadrillion Btu in 2005. In the residential sector, growth was from 1.1 to 14.6 quadrillion Btu. In 2005, 28.6 quadrillion Btu of the 39.6 quadrillion total electricity consumption was devoted to the commercial and residential sectors.

Most of the remaining electricity usage was in the industrial sector. Electricity consumption grew from 2.1 quadrillion Btu in 1949 to about 10.9 quadrillion in 2005. Direct consumption of coal went from about 5.4 to about 2 quadrillion Btu. Natural gas consumption grew from 3.2 quadrillion to 7.4 quadrillion Btu. Petroleum consumption grew from 3.5 to 9.7 quadrillion Btu.

The transportation sector is dominated by fossil fuel consumption. Over the 1949 – 2005 period, between 98% and 99.9% of consumption comes from fossil fuels. The primary fuel is petroleum. In 1949 about 77% of total transportation energy consumption was from petroleum. By 1957, the petroleum share was 96%. From 1957 through 2005, the petroleum share stayed in the 96% to 98% range. Overall, most substitution of fossil fuels within the sectors of the economy was from direct consumption of coal, natural gas, or petroleum to indirect consumption that resulted from the emergence and growth of the electricity sector.

After estimating the demand models, we estimate the contributions of each fossil fuel to human CO_2 emissions and compare the forecast ability of this model to a model that estimates the contributions of total fossil fuel energy to CO_2 emissions. We find that the model that separates the fossil fuels by type outperforms the aggregate fossil fuel energy model in terms of ability to forecast CO_2 emissions in a non-systematic fashion. We attribute the improved forecast ability of the disaggregated model to its ability to disentangle the change in energy mix that occurs due to differing short run and long run responses to changes in real GDP.

Data

Data are from the Energy Information Administration's Annual Energy Review 2006[1]. The variables are consumption of coal, crude oil and natural gas measured in quadrillions of British Thermal Units (Btu), prices of coal, crude oil and natural gas, measured in 2000 GDP chain price deflated dollars per million Btu, carbon dioxide emissions from energy, measured in millions of metric tons, and real GDP in billions of 2000 chain price indexed dollars. We also aggregate total fossil energy consumption and a fossil energy price index. Total fossil energy consumption is the summed consumption of coal, crude oil and natural gas. The price index is derived from the first principal component of the prices of coal, crude oil, and natural gas. These data are for the years 1949-2005.

We create the price index with principal component analysis rather than quantity weighting to avoid endogeneity that may result from having total fossil energy consumption on both sides of a regression equation. If we weight prices of each energy source by the share of total energy consumption, the denominator of each weight is total energy consumption. If we regress this price index on total energy consumption, then total energy consumption is necessarily included on both sides of the equation and the price index is correlated with the regression disturbance.

Principal components analysis has been used to create price indexes since at least 1970 (see Doll and Chin (1970)). The procedure reduces the information from several variables into fewer variables. We find the translation of the three dimensional Cartesian plane that minimizes the variance among these prices. Each of the three resulting vectors is an orthogonal projection of the covariance matrix. The original price vectors are now a linear combination of the new orthogonal vectors. These new vectors are determined by deriving the three eigenvalues for the covariance matrix of prices. The largest eigenvalue is the first principal component. In our case, the first principal component accounts for 88% of the total variance in the system. The eigenvector associated with the first principal component is (-.098031, -.573934, -.813012), where the values are (coal price, gas price, oil price). This eigenvector is unique to a scalar multiple, which we chose as -0.6734 (This multiple is

$\frac{1}{-0.098031 + -.573934 + -.813012}$) giving the eigenvector (.066, .386, .548) which serves as our weights for each price. Thus, Price Index$_t$ = .066*coal price$_t$ + .386*gas price$_t$ + .548*oil price$_t$. The first principal component accounts for 88% of the variation among the original prices and yields a price index that is uncorrelated with the regression disturbance.

There are 14 variables used in the regressions: 4 prices in natural logarithms, 4 consumption measures in natural logarithms and 4 consumption measures in quadrillions Btu, carbon dioxide emissions in millions of metric tons, and real GDP measured in natural logarithms. If these series are non-stationary, spurious inferences may result from regression analysis. We use Augmented Dickey-Fuller tests for unit roots. Only the natural logarithm on natural gas consumption rejects the hypothesis of a unit root in levels. All other time series are stationary in first differences.

Econometric Models

We have four models of energy demand. The dependent variable is the consumption of coal, natural gas, crude oil, or total fossil fuel energy, measured in natural logarithms and the regressors are the prices of the different fuels and GDP expressed in natural logarithms. Equations (1) and (2) formally express the models.

4.1 $$\ln F_{it}^{*} = \beta_{i0} + \sum_{j} \beta_{ij} \ln P_{jt} + \beta_{i4} \ln GDP_{t} + \beta_{i5} Time$$

4.2 $$\ln TotalEnergy_{t}^{*} = \alpha_{o} + \alpha_{1} \ln PriceIndex_{t} + \alpha_{2} \ln GDP_{t} + \alpha_{3} time$$

Equation 4.1 states that the long run demand for fuel i = crude oil, natural gas, or coal depends on the fuel prices and real GDP. β_{ij} is the cross-price elasticity of demand for fuel i with respect to price j. When i=j, this is the own price elasticity of demand. β_{i4} is the income elasticity of demand. The coefficients on time measure changes in consumption over time, presumably due to changes in technology and preferences.

Equation 4.2 models total energy demand as a function of the energy price index and GDP. Price and income elasticities are α_1 and α_2.

Because adjustments in fuel consumption may not occur annually, we adopt a partial adjustment model, described in equation 4.3.

4.3 $$(1-\lambda)(Y_{t}^{*} - Y_{t-1}) + \varepsilon_{t} = (Y_{t} - Y_{t-1}),$$

where $Y_{t}^{*} = f(\vec{X})$ describes the long run equilibrium relationship between Y and the covariates $\vec{X}$. $(1-\lambda)$ is the fraction of adjustment toward the long run equilibrium that occurs in one year. This framework yields regression equations 4.4 and 4.5.

4.4 $$\ln F_{it} = \beta_{i0}^{+} + \sum_{j} \beta_{ij}^{+} \ln P_{jt} + \beta_{i4}^{+} \ln GDP_{t} + \beta_{i5}^{+} Time + \lambda_{i} \ln F_{it-1} + \varepsilon_{it}$$

4.5 $$\ln TotalEnergy_{t} = \alpha_{0}^{+} + \alpha_{1}^{+} \ln PriceIndex_{t} + \alpha_{2}^{+} \ln GDP_{t} + \alpha_{3}^{+} Time + \lambda_{TE} \ln TotalEnergy_{t-1} + \varepsilon_{TE,t}$$

The "+" superscript represents the short run coefficients. For example, α_1^+ is the short run price elasticity of demand for total energy consumption. The long run price elasticity is $\alpha_1 = \dfrac{\alpha_1^+}{1-\lambda_{TE}}$. Wald tests are performed to determine the statistical significance of the nonlinear long run coefficient estimates. All regressions are estimated with variables in first differences, denoted by D(.). Estimating the equations in first differences has two effects on equations 4.4 and 4.5. First, it eliminates the constant terms from these regression equations. Second, since the difference between time periods is one year, the value of the first differenced time variable is always one. The net effect of these changes is that the constant term in first differenced regressions can be interpreted as changes in consumption over time due to changes in technology and preferences.

Tables 4.1 through 4.4 list the regression results for the coal, natural gas, crude oil, and total fossil energy demand. Breusch-Godfrey tests fail to reject the hypothesis of serially independent disturbances.

The coal demand equation provides little information about the economic influences on coal consumption. According to the model, $1-\lambda_c$, the annual adjustment to long run equilibrium, is 0.965 and not statistically different from one, implying full adjustment in one year. The long run and short run income elasticities are about 1.39 and statistically different from 0 but not from 1. The short run and long run price and cross-price elasticities are not statistically significant.

Table 4.1. Coal Demand 1949-2005

Dependent Variable: DLNCOAL
Method: Least Squares

Sample (adjusted): 1951 2005
Included observations: 55 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.036566	0.010722	-3.410286	0.0013
LNPriceCOAL	0.008919	0.063896	0.139580	0.8896
DLNPriceGAS	0.006896	0.041485	0.166218	0.8687
DLNPriceOIL	0.037586	0.033302	1.128647	0.2645
DLNGDP	1.393685	0.261110	5.337528	0.0000
DLNCOAL(-1)	0.035897	0.114578	0.313301	0.7554

R-squared	0.389538	Mean dependent var	0.011209
Adjusted R-squared	0.327246	S.D. dependent var	0.048229
S.E. of regression	0.039558	Akaike info criterion	-3.519416
Sum squared resid	0.076678	Schwarz criterion	-3.300434
Log likelihood	102.7839	F-statistic	6.253407
Durbin-Watson stat	1.772969	Prob(F-statistic)	0.000147

The gas demand model is more informative. $1-\lambda_g$, the one-year adjustment in gas consumption, is 0.50 and statistically different from both zero and one. This indicates that about 50% of the adjustment to long run equilibrium occurs within the first year. The estimated cross-price elasticity with respect to coal prices is -0.069 but not significant. A priori, we would expect this coefficient estimate to be positive, indicating coal and gas are substitutes. The lack of statistical significance indicates that there is no detectable relationship between gas demand and coal prices.

Both the short run and long run elasticities of gas demand with respect to oil prices are statistically significant. The short run estimate indicates a 1% increase in oil prices causes a 0.088% increase in the demand for natural gas in the short run. In the long run, the cross-price elasticity is .178, indicating that a permanent 1% increase in oil prices leads to a .177% increase in natural gas consumption. The sign of the coefficient estimates indicates that oil is a substitute for natural gas, consistent with a priori expectations.

Table 4.2. Gas Demand 1949-2005

Dependent Variable: DLNGAS
Method: Least Squares

Sample (adjusted): 1951 2005
Included observations: 55 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.010015	0.008588	-1.166142	0.2492
DLNPriceCOAL	-0.068647	0.050948	-1.347377	0.1841
DLNPriceGAS	-0.110000	0.033290	-3.304267	0.0018
DLNPriceOIL	0.087979	0.026735	3.290736	0.0019
DLNGDP	0.725548	0.207019	3.504745	0.0010
DLNGAS(-1)	0.504628	0.088084	5.728945	0.0000

R-squared	0.580733	Mean dependent var	0.024241
Adjusted R-squared	0.537950	S.D. dependent var	0.046705
S.E. of regression	0.031747	Akaike info criterion	-3.959355
Sum squared resid	0.049386	Schwarz criterion	-3.740373
Log likelihood	114.8823	F-statistic	13.57410
Durbin-Watson stat	2.141679	Prob(F-statistic)	0.000000

The own price elasticities are correctly signed and statistically different from zero. The short run estimate is -0.11 while the long run estimate is -0.22, indicating that a 1% increase in gas prices causes a 0.11% short run and a 0.22% long run demand reduction. The estimated short run income elasticity for gas consumption is 0.73 and the long run elasticity is 1.46. Both are statistically different from zero but not from one.

Oil prices and GDP are the only statistically significant variables in the oil demand model. A priori, one would expect both coal and gas to be substitutes for oil.

However, oil consumption occurs chiefly in the transportation sector, where neither coal nor natural gas is a feasible alternative. In 2006, the U.S. consumed 39.758 quadrillion Btu of petroleum products; 27.248 quadrillion Btu or 68.5% of that consumption was in the transportation sector. Total consumption of all energy in the transportation sector was 28.4 quadrillion Btu[2].

Table 4.3. Oil Demand 1949 - 2005

Dependent Variable: DLNOIL
Method: Least Squares

Sample (adjusted): 1951 2005
Included observations: 55 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.014212	0.005603	-2.536534	0.0144
DLNPriceCOAL	-0.020197	0.036384	-0.555113	0.5813
DLNPriceGAS	0.032345	0.022611	1.430538	0.1589
DLNPriceOIL	-0.059060	0.018020	-3.277548	0.0019
DLNGDP	0.742585	0.152405	4.872463	0.0000
DLNOIL(-1)	0.412468	0.090674	4.548926	0.0000

R-squared	0.647972	Mean dependent var	0.020199
Adjusted R-squared	0.612051	S.D. dependent var	0.034341
S.E. of regression	0.021389	Akaike info criterion	-4.749174
Sum squared resid	0.022418	Schwarz criterion	-4.530192
Log likelihood	136.6023	F-statistic	18.03871
Durbin-Watson stat	1.597087	Prob(F-statistic)	0.000000

We estimate 1-λ_o as 0.59 and statistically different from both one and zero. This indicates that 59% of the adjustment toward long run equilibrium occurs within one year. The short run price elasticity is -0.059 and the long run elasticity about -0.10. The short run income elasticity of demand is 0.74 and statistically different from zero and one. The long run income elasticity estimate of 1.25 is statistically larger than one.

Table 4.4. Total Energy Demand 1949 - 2005

Dependent Variable: D(LNTOTAL)
Method: Least Squares

Sample (adjusted): 1951 2005
Included observations: 55 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.019094	0.005006	-3.814292	0.0004
DLNPENERGY	-0.009884	0.015369	-0.643127	0.5230
DLNGDP	1.024405	0.122044	8.393737	0.0000
D(LNTOTAL(-1))	0.175938	0.083497	2.107130	0.0400

R-squared	0.619693	Mean dependent var	0.018176
Adjusted R-squared	0.597322	S.D. dependent var	0.030387
S.E. of regression	0.019282	Akaike info criterion	-4.989300
Sum squared resid	0.018962	Schwarz criterion	-4.843312
Log likelihood	141.2057	F-statistic	27.70076
Durbin-Watson stat	1.629751	Prob(F-statistic)	0.000000

Consider total energy shown in Table 4.4. The principal component price index statistically insignificant probably because aggregating prices of the different fuels causes a loss of information regarding interfuel substitution that is important in terms of total energy demand. Overall, this model fits the data well. The R^2 value for the regression with variables in first differences is 0.62. The one-year adjustment estimate is 0.825 and different from both zero and one. The short run income elasticity of 1.02 is different from zero but not one. The long run income elasticity estimate of 1.24 is likewise not statistically different from one.

Modeling Carbon Dioxide Emissions

Recent work has examined the role of energy consumption on carbon dioxide emissions. Schmalansee, et al (1998) extended the panel data set of Holtz-Eakin and Selden (1995) to include more countries over the 1950 to 1990 period. They specify a model of CO_2 emissions per capita in country j at time t as function of country specific effects, year specific effects, and lnGDP per capita in country j at time t. We believe such a specification can lead to erroneous conclusions regarding future CO_2 emissions. First, per capita CO_2 emissions often trend differently from total emissions. In the United States, per capita emissions peaked in 1972 and declined until 1985. Per capita emissions have risen annually since. Total emissions increased throughout the years 1949-2005. Second, our simple energy demand models show that the different types of fossil fuels have different responses to GDP so that over time the energy mix within a country changes in response to changes in GDP.

To illustrate the significance of the changing energy mix, we compare forecasts of CO_2 emissions based on total energy consumption to forecasts based on consumption of each fossil energy source. We do this by estimating equations 4.4 and 4.5 using only data from 1949-2000. The data for years 2001 – 2005 are then used for computing out-of-sample forecasts for each type of energy.

The results of the coal, natural gas, oil, and total energy forecasts are reported in tables 4.5A – 4.8B. Table 4.5A reports the estimation results for the coal model of equation (4). Table 4.5B reports the out-of-sample forecast for coal consumption based on these coefficient estimates and compares it to actual coal consumption. Tables 4.6A and 4.6B do the same for natural gas and Tables 4.7A and 4.7B do this for crude oil. Table 4.8A reports the sub-sample coefficient estimates for total energy consumption and 4.8B compares the actual and predicted total energy consumption.

The sub-sample coefficient estimates differ little from full-sample estimates. For the coal demand equations, full adjustment still occurs within one year, the own price and cross price elasticities are insignificant, and the income elasticity is approximately the same. The largest forecast error is about 2% with a mean absolute percent error of about 1.2%

Table 4.5A. Coal Demand 1949-2000 Sub-Sample

Dependent Variable: DLNCOAL
Method: Least Squares
Date: 04/15/08 Time: 11:16
Sample (adjusted): 1951 2000
Included observations: 50 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.038249	0.011960	-3.197938	0.0026
DLNPCOAL	0.016006	0.069483	0.230356	0.8189
DLNPGAS	0.012792	0.053613	0.238596	0.8125
DLNPOIL	0.037492	0.037498	0.999846	0.3229
DLNGDP	1.433750	0.282472	5.075725	0.0000
DLNCOAL(-1)	0.045129	0.122237	0.369192	0.7138

R-squared	0.387367	Mean dependent var	0.012129
Adjusted R-squared	0.317750	S.D. dependent var	0.050221
S.E. of regression	0.041482	Akaike info criterion	-3.414953
Sum squared resid	0.075713	Schwarz criterion	-3.185510
Log likelihood	91.37381	F-statistic	5.564236
Durbin-Watson stat	1.781635	Prob(F-statistic)	0.000468

Table 4.5B. Coal Forecasts 2001 - 2005

Year	Actual Coal Consumption (quadrillion btu)	Forecasted Coal Consumption (quadrillion btu)	Difference between Actual and Forecasted Consumption (quadrillion btu)	Absolute Value of the Percent Error of the Forecast
2001	21.944	21.933	0.010	0.0
2002	21.965	21.524	0.441	2.0
2003	22.371	22.203	0.169	0.8
2004	22.604	23.068	-0.464	2.1
2005	22.874	23.195	-0.321	1.4

The natural gas sub-sample estimation is also very similar to that of the full sample. There is still no detectable relationship with coal, and oil is a substitute. The estimated own price elasticity drops from -0.11 to -0.08 in the short run and from -0.22 to -0.18 in the long run. The short run and long run income elasticities are approximately the same, as is the one year adjustment towards long run equilibrium.

The out of sample forecast over-predicts natural gas consumption in 4 of the 5 years. The forecast error is less than 4% in all years with a mean forecast error of about 2.3%.

In the oil demand equation, the one-year adjustment estimate is 59% for the full sample and 57% for the sub-sample. Oil prices and GDP are still the only significant regressors. The short run and long run price and income elasticities are comparable across the samples. The short run estimate of the natural gas price is marginally significant indicating that natural gas is possibly a substitute for oil.

Table 4.6A. Gas Demand 1949-2000 Sub-Sample

Dependent Variable: DLNGAS
Method: Least Squares

Sample (adjusted): 1951 2000
Included observations: 50 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.010279	0.009493	-1.082830	0.2848
DLNPriceCOAL	-0.061966	0.054433	-1.138393	0.2611
DLNPriceGAS	-0.082166	0.042467	-1.934817	0.0595
DLNPriceOIL	0.076228	0.029635	2.572252	0.0136
DLNGDP	0.722344	0.219812	3.286194	0.0020
DLNGAS(-1)	0.517355	0.094533	5.472741	0.0000

R-squared	0.561509	Mean dependent var	0.027762
Adjusted R-squared	0.511680	S.D. dependent var	0.046868
S.E. of regression	0.032751	Akaike info criterion	-3.887597
Sum squared resid	0.047196	Schwarz criterion	-3.658155
Log likelihood	103.1899	F-statistic	11.26882
Durbin-Watson stat	2.118348	Prob(F-statistic)	0.000000

Table 4.6B. Gas Forecast 2001-2005

Year	Actual Gas Consumption (quadrillion btu)	Forecasted Gas Consumption (quadrillion btu)	Difference between Actual and Forecasted Consumption (quadrillion btu)	Absolute Value of the Percent Error of the Forecast
2001	22.906	23.718	-0.812	3.5
2002	23.628	23.033	0.595	2.5
2003	22.967	23.624	-0.657	2.9
2004	23.036	23.221	-0.185	0.8
2005	22.640	23.076	-0.436	1.9

Table 4.7A. Oil Demand 1949-2000 Sub-Sample

Dependent Variable: DLNOIL
Method: Least Squares

Sample (adjusted): 1951 2000
Included observations: 50 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.015377	0.005963	-2.578468	0.0133
DLNPriceCOAL	-0.020650	0.037973	-0.543800	0.5893
DLNPriceGAS	0.049472	0.028124	1.759037	0.0855
DLNPriceOIL	-0.070013	0.019502	-3.589967	0.0008
DLNGDP	0.726393	0.156862	4.630769	0.0000
DLNOIL(-1)	0.438896	0.092476	4.746069	0.0000

R-squared	0.671168	Mean dependent var	0.021184
Adjusted R-squared	0.633801	S.D. dependent var	0.035535
S.E. of regression	0.021504	Akaike info criterion	-4.729003
Sum squared resid	0.020346	Schwarz criterion	-4.499560
Log likelihood	124.2251	F-statistic	17.96141
Durbin-Watson stat	1.594807	Prob(F-statistic)	0.000000

Table 4.7B. Oil Forecast 2001 - 2005

Year	Actual Oil Consumption (quadrillion btu)	Forecasted Oil Consumption (quadrillion btu)	Difference between Actual and Forecasted Consumption (quadrillion btu)	Absolute Value of the Percent Error of the Forecast
2001	38.333	38.938	-0.605	1.6
2002	38.401	37.502	0.899	2.3
2003	39.047	38.996	0.051	0.1
2004	40.594	39.216	1.378	3.4
2005	40.441	41.238	-0.797	2.0

Table 4.8A. Total Energy Demand 1949 - 2000

Dependent Variable: D(LNTOTAL)
Method: Least Squares

Sample (adjusted): 1951 2000
Included observations: 50 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.018925	0.005443	-3.476997	0.0011
DLNPENERGY	-0.007034	0.017042	-0.412729	0.6817
DLNGDP	1.017833	0.129032	7.888243	0.0000
D(LNTOTAL(-1))	0.191678	0.087639	2.187136	0.0339

R-squared	0.617932	Mean dependent var	0.019761
Adjusted R-squared	0.593014	S.D. dependent var	0.031122
S.E. of regression	0.019854	Akaike info criterion	-4.924179
Sum squared resid	0.018133	Schwarz criterion	-4.771217
Log likelihood	127.1045	F-statistic	24.79912
Durbin-Watson stat	1.592748	Prob(F-statistic)	0.000000

Table 4.8B. Forecast of Total Energy Consumption 2001- 2005

Year	Actual Energy Consumption (quadrillion btu)	Forecasted Energy Consumption (quadrillion btu)	Difference between Actual and Forecasted Consumption (quadrillion btu)	Absolute Value of the Percent Error of the Forecast
2001	83.182	86.415	-3.233	3.9
2002	83.994	83.367	0.627	0.7
2003	84.386	84.980	-0.594	0.7
2004	86.233	85.242	0.991	1.1
2005	85.955	87.539	-1.584	1.8

Forecasts are accurate. The largest forecast error is 3.4% with a mean absolute percent error of 1.9%. The forecast errors are not systematic either. There are two years of over-predicting and three years of under-predicting oil consumption.

The total energy demand model follows suit with the rest of the demand models in terms of robustness and forecast accuracy. The one-year adjustment for the full sample is 82.5% compared to 81% for the sub-sample. The composite price elasticity is not statistically significant for either model. The short run and long run income elasticities are both statistically different from zero but not from one and both pairs of income elasticities are virtually the same across sample periods.

The largest forecast error is 4% with a mean absolute percent error of 1.64%. The errors are not systematic. Three years over-predict and two years under-predict total energy consumption.

The second step is to estimate carbon dioxide emissions from each fuel using equation 4.6 and for aggregate fuel using equation 4.7.

4.6 $$D(CO_2)_t = constant + \beta_c D(Coal)_t + \beta_G D(Gas)_t + \beta_O D(Oil)_t + \varepsilon_t$$

4.7 $$D(CO_2)_t = constant + \beta_{TE} D(TotalEnergy)_t + v_t$$

In principle, there is no need to estimate the coefficients in these models. The EIA provides CO_2 coefficients for each type of fossil fuel[3]. However, the CO_2 coefficients differ by grade of coal, refined petroleum product, or whether natural gas is flared or shipped by pipeline. These coefficients also change over time. Our data provide quantities of aggregated coal, petroleum, and natural gas consumption. We do not have the means to disaggregate consumption to the levels necessary to use the true coefficient values. The regressions provide a close approximation to the actual coefficients. The equations are first differenced because Augmented Dickey Fuller tests could not reject a unit root for the time series of each variable. The results of equation 4.6 are shown in table 4.9.

The model fits remarkably well. The R^2 value is 0.995 and Durbin-Watson and Breusch-Godfrey tests fail to reject the null hypotheses of no serial correlation. The coefficient estimates are highly significant for all three fuels: $\beta_c = 89.00$, indicating that a 1 quadrillion Btu increase in coal consumption increases human CO_2 emissions by 89 million metric tons. The EIA coefficients for coal combustion are approximately 93, which falls within the 95% confidence interval of our estimate. The β_G estimate indicates that a 1 quadrillion Btu increase in natural gas consumption increases human CO_2 emissions by 51.56 million metric tons. The true coefficient for flared gas is 54.71 and the coefficient for piped gas is 52.79. Both coefficients are within the 95% confidence interval of our estimate. The β_O estimate implies a one quadrillion Btu increase in oil consumption leads to a 67.89 million metric ton increase in CO_2 emissions. The EIA coefficients for unrefined crude oil range from 72.31 in 1980 to 73.79 in 2005. The coefficients for refined products range from 62.09 for liquid propane gas to 101.1 for petroleum coke. The most common refined product is motor gasoline with a coefficient of 70.46, nearly identical to our estimate of 67.89.

The results of equation 4.7 are given in Table 4.10. The estimated coefficient on total energy indicates that a one quadrillion Btu increase in fossil energy consumption yields a 66.86 million metric ton increase in CO_2 emissions.

Table 4.9. Carbon Dioxide Emissions Equation for Fossil Fuels, 1949 - 2000

Dependent Variable: D(CO2)
Method: Least Squares
Date: 07/30/07 Time: 13:01
Sample (adjusted): 1950 2000
Included observations: 51 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.800213	1.462011	-1.231327	0.2243
D(COAL)	88.99935	2.089591	42.59175	0.0000
D(GAS)	51.55828	1.903185	27.09052	0.0000
D(OIL)	67.89416	1.317378	51.53734	0.0000

R-squared	0.994809	Mean dependent var	71.10392
Adjusted R-squared	0.994477	S.D. dependent var	118.3398
S.E. of regression	8.794377	Akaike info criterion	7.261287
Sum squared resid	3635.030	Schwarz criterion	7.412803
Log likelihood	-181.1628	F-statistic	3002.203
Durbin-Watson stat	2.379682	Prob(F-statistic)	0.000000

Fossil fuel values are in quadrillion btu. Carbon Dioxide emissions are in millions of metric tons. The coefficient estimate of 88.999 for coal indicates that a one quadrillion btu increase in coal consumption causes an 88.999 million metric ton increase in CO_2 emissions.

The results of equations 4.6 and 4.7 vary greatly. Equation 4.6 shows that fossil fuel sources vary greatly in their CO_2 content. As expected, coal emits the most CO_2 per Btu and natural gas the least (Jorgenson and Wilcoxen, 1992). Equation 4.7 indicates that CO_2 emissions based on aggregated energy usage are slightly less than that of a Btu of oil.

Table 4.10. Carbon Dioxide Emissions based on Total Energy 1949-2000

Dependent Variable: D(CO2)
Method: Least Squares

Sample (adjusted): 1950 2000
Included observations: 51 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.264553	3.037408	-0.745554	0.4595
D(TOTAL/1000000)	66.86207	1.481557	45.12958	0.0000

R-squared	0.976506	Mean dependent var	71.10392
Adjusted R-squared	0.976027	S.D. dependent var	118.3398
S.E. of regression	18.32282	Akaike info criterion	8.692597
Sum squared resid	16450.55	Schwarz criterion	8.768355
Log likelihood	-219.6612	F-statistic	2036.679
Durbin-Watson stat	1.401546	Prob(F-statistic)	0.000000

Energy Consumption values are in quadrillion btu. Carbon Dioxide emissions are in millions of metric tons. The coefficient estimate of 66.86 for coal indicates that a one quadrillion btu increase in coal consumption causes an 88.999 million metric ton increase in CO_2 emissions.

Next, we use the coefficient estimates from equations 4.6 and 4.7 to forecast CO_2 emissions from 2001 to 2005. The results, reported in Tables 4.11 for disaggregated energy, are remarkably accurate. Using forecasted values of the three fossil fuels in a forecast of CO_2 emissions yielded errors that were less than 2% for all years. The mean absolute percentage error for the 5 year out of sample period is 1.26%.

Table 4.12 reports forecasts using equation 4.7. This equation also forecasts well but always under-predicts CO_2 emissions and the mean absolute forecast error is larger than that of the model based on individual energy demand forecasts.

Table 4.11. Carbon Dioxide Emissions Forecast 2001 – 2005 based on Equation 4.6

Year	Actual CO2 Emissions (million metric tons)	Forecasted CO2 Emissions (million metric tons)	Difference between Actual and Forecasted Consumption (million metric tons)	Absolute Value of the Percent Error of the Forecast
2001	5709.8	5766.8	-57.0	1.0
2002	5752.2	5599.0	153.2	2.7
2003	5800.5	5789.9	10.6	0.2
2004	5923.2	5861.3	61.9	1.0
2005	5945.3	6002.4	-57.1	1.0

Table 4.12. CO_2 Forecast from Equation 4.7

Year	Actual CO2 Emissions (million metric tons)	Forecasted CO2 Emissions (million metric tons)	Difference between Actual and Forecasted Consumption (million metric tons)	Absolute Value of the Percent Error of the Forecast
2001	5709.8	5498.5	211.3	3.7
2002	5752.2	5552.2	200.0	3.5
2003	5800.5	5578.1	222.4	3.8
2004	5923.2	5700.2	223.0	3.8
2005	5945.3	5681.8	263.5	4.4

Conclusion

The conclusion is clear. Despite the fact that the total energy consumption model forecasted more accurately and less systematically than did any of the individual fossil fuel models, the CO_2 emissions forecast based on the total energy forecast was more systematic and less accurate. The reason for this disparity is that the aggregated fossil energy demand model fails to account for the fuel mix which is affected by GDP. As a result, it fails to account for the changes in CO_2 that correspond with the changing fuel mix. The aggregate model cannot distinguish the early years of the sample, where coal and petroleum consumption each accounted for about 41% of the total fossil energy usage, from the contemporary period, where natural gas consumption is comparable to that of coal at about 27% of total consumption.

The inability to distinguish the energy mix implies that the aggregate model is incapable of distinguishing changes is CO_2 emissions that result from the change in energy usage patterns. Natural gas emits about 42% less CO_2 per Btu than coal. If a model cannot distinguish between coal and natural gas consumption, it will clearly provide less accurate forecasts than a model that distinguishes between the usage patterns.

Notes

The following information appeared as superscripted notes in the text of this chapter.

1. Data are available upon written request.
2. Figures from the transportation sector come from Table 2.1e of the EIA Annual Energy Review 2006.
3. Perry Lindstrom of the Energy Information Administration provided the carbon dioxide coefficients cited.

CHAPTER V

CONCLUSION

Chapter II analyzes the refining industry in Texas, Oklahoma, and Louisiana. The key questions being addressed are why the scale of operation increased and the impact this increase had on the demand for production workers. The key conclusions are that technological progress, increasing relative wages and the cost of compliance with environmental regulation caused refineries to increase the scale of plants and reduced both the number of plants in operation and the quantity of labor used to produce a given volume of refined products. Although oil companies such as ExxonMobil recently reported record profits, there is no evidence to suggest this is the result of collusive behavior in the refining industry. These profits were likely the result of vertical integration in these firms. The most profitable section of these companies was oil production. The price of oil increased as a result of increasing world demand and this price increase was incorporated as a cost of production in the refining industry.

The recent high prices of oil have motivated many policy proposals within the U.S. Congress. These proposals advocate increasing access for exploration and development on federal lands, reducing the royalties paid to the government for oil produced on public property, providing subsidies for production, or tax breaks for exploration and development. Chapter III analyzes the potential effectiveness of these proposals by modeling U.S. oil production from 1982 to 2003. The major finding of this chapter is that both the price elasticity of supply and the output elasticity with respect to reserves are about 0.23 in the long run. The implication of this result is that proposals designed to stimulate domestic production will have little effect.

We also simulate production between 2005 and 2030 using various increases in prices and reserves. The scenario that results in the highest annual production assumes that real prices double over this period and that reserves increase by 50% after 2013 due to increased access to the Alaska National Wildlife Refuge. Under this best case for production scenario, production in 2030 is approximately the same as that in 2005. Over this period, the Energy Information Administration projects that domestic consumption will grow by 2.5 percent annually. This implies that if real prices double and reserves increase by 50%, production will remain at about its 2005 level while consumption nearly doubles. The oil trade deficit will increase substantially and the already mature domestic oilfields will be further exhausted.

Chapter IV analyzes the demand for fossil fuel energy. The importance of fossil energy cannot be overstated. 85 percent of all U.S. energy consumption comes from fossil fuels. An interesting finding is that there is virtually no substitution between the fuels. It is more appropriate to refer to three separate markets than it is to refer to a domestic energy market. Over 90 percent of all coal is used for electricity production and 51 percent of all electricity is generated from coal fired plants. Natural gas accounts for about 16 percent and oil about 1.6 percent of electricity generation but the technologies for generating electricity from the different fuels are vastly different and substantial capital investment is necessary to change the production mix. Natural gas is used primarily in the residential heating sector. Oil and coal are not easily interchangeable for home heating and substantial capital investment is required to change the fuel mix. The same result holds for oil. It accounts for nearly 40 percent of all energy consumption. About 69% of oil consumption was in the transportation industry where oil accounted for 98 percent of all energy consumed. Coal and natural gas are not feasible substitutes for oil in transportation.

The lack of substitutability documented above is confirmed by the coefficient estimates in the demand equations. The only statistically significant cross-price elasticity indicated that crude oil was a substitute for natural gas. No other cross price elasticity was significant. The energy models also indicate that demand for all products is driven more by GDP than by prices and that each fuel has a substantially different GDP elasticity. Coal is used to generate electricity which is the primary energy for industrial and commercial uses. It has the largest response to a change in GDP with an elasticity of about 1.5 in both the short run and the long run. The short and long run GDP elasticities for gas are 0.72 and 1.45 and those for oil are 0.74 and 1.25. No fuel has a long run price elasticity of demand greater than 0.22.

The importance of these results is illustrated through forecasts of CO_2 emissions. We estimate that coal emits about 90 million metric tons of CO_2 per quadrillion Btu; oil about 67.9 and gas about 51.6. Over time the energy mix changes as a result of differing responses to GDP. Forecasts of CO_2 emissions generated from forecasts of each fossil fuel are more accurate than CO_2 forecasts generated from a forecast of aggregate fossil fuel consumption. This is counter to conventional wisdom. Generally, parsimonious models provide more accurate forecasts. The reason that a model with more explanatory variables is preferable for energy based CO_2 emissions is the disparity in emissions between the different fuels. Gas emits about 57% of the CO_2 of coal. If energy consumption increases as a result of increased gas consumption the CO_2 emissions will be different than if total energy consumption changes are driven by increases in coal consumption. An aggregate model of energy consumption cannot distinguish between these events. The ability of the disaggregated model to distinguish these events allows it to provide more accurate forecasts.

REFERENCES

Anderson, Theodore and Cheng Hsiao. 1981. "Estimation of Dynamic Models with Error Components." *Journal of the American Statistical Association,* 76: 869-881.

Arellano, Manuel and Stephen Bond. 1991, "Some Tests for Specification of Panel Data: Monte Carlo Evidence and an Application to Equations." *Review of Economic Studies,* 58: 127-134.

Baltagi, Badi. 2001. *Econometric Analysis of Panel Data*, 2nd edition, Chichester U.K.: John Wiley & Sons.

Baltagi, Badi, James Griffin, and Weiwen Xiong. 2000. "To Pool or Not to Pool: Homogeneous Versus Heterogeneous Estimators Applied to Cigarette Demand." *Review of Economics and Statistics*, 82: 117-126.

Bruno, Giovanni. 2005a. "Approximating the Bias of the LSDV Estimator for Dynamic Unbalanced Panel Data Models." *Economics Letters*, 87(3), 361-366.

Bruno, Giovanni. 2005b. "Estimation and Inference in Dynamic Unbalanced Panel-Data Models with a Small Number of Individuals." *Stata Journal*, 5(4), 473-500.

Cleveland, Cutler. 1991. "Physical and Economic Aspects of Resource Quality: The Cost of Oil Supply in the Lower 48 United States, 1938-1988." *Resources and Energy,* 13: 163-188.

Cleveland, Cutler and Robert Kaufmann. 1991. "Forecasting Ultimate Oil Recovery and Its Rate of Production: Incorporating Economic Forces into the Models of M. King Hubbert." *The Energy Journal,* 12(2): 17-46.

Doll, John and Sean Chin. 1970. "A Use of Principal Components in Price Analysis." *American Journal of Agricultural Economics,* 39(4): 975-1001.

Energy Information Administration (EIA). 2000. *Potential Oil Production from the Coastal Plain of the Arctic National Wildlife Refuge: Updated assessment,* Washington DC.

Energy Information Administration (EIA). 2006. *Documentation for Emissions of Greenhouse Gases in the United States 2004,* Washington DC.

Energy Information Administration (EIA). 2007. *Annual Energy Review 2006,* Washington DC.

Griffin, James. 1985. “OPEC Behavior: A Test of Alternative Hypotheses.” *American Economic Review,* 75: 954-963.

Hamilton, James. 1983. “Oil and the Macroeconomy Since WWII.” *Journal of Political Economy,* 91(2): 228-248.

Hamilton, James and Ana Herrera. 2004. “Oil Shocks and Aggregate Macroeconomic Behavior: The Role of Monetary Policy.” *Journal of Money Credit and Banking,* 36(2): 265-286.

Hibdon, James and Michael Mueller. 1990. “Economies of Scale in Petroleum Refining, 1947-1984: A Survivor Principle-Time Series Analysis.” *Review of Industrial Organization,* 5(3): 25-43.

Holtz-Eakin, Douglas and Thomas Selden. 1995. “Stoking the Fires? CO_2 Emissions and Economic Growth.” *Journal of Public Economics,* 57: 85-101.

Hotelling, Harold. 1931. “The Economics of Exhaustible Resources.” *Journal of Political Economy,* 39: 137-175.

Hu, Ling. 2008. “A Simple Panel Cointegration Test and Estimation in I(0)/I(1) Mixed Panel Data.” unpublished working paper, Department of Economics, Ohio State University, Columbus.

Hubbert, M. King. 1962. Energy Resources. A Report to the Committee on Natural Resources, Publication 1000-D. Technical report, Washington, DC: National Academy of Sciences.

Hubbert, M. **King**. 1967. “Degree of Advancement of Petroleum Exploration in the United States.” *American Association of Petroleum Geologists Bulletin,* 51: 2207-2227.

Hubbert, M. King. 1982. “Techniques of Prediction as Applied to the Production of Oil and Gas.” In *Oil and Gas Supply Modeling*, ed. S. Glass, 16-141. Washington, D.C.: National Bureau of Standards, Special Publication 631.

Hubbert, M. King.. 1956. “Nuclear Energy and the Fossil Fuels.” In *Drilling and Production Practices*, pp. 7-25. Washington, DC: American Petroleum Institute.

Intergovernmental Panel on Climate Change (IPCC). 2007. "Summary for Policy Makers." In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report to the Intergovernmental Panel on Climate Change.* eds. Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller. Cambridge: Cambridge University Press. 1-18.

Im, Kyung, Hashem Pesaran, and Yongcheol Shin. 2003. "Testing for Unit Roots in Heterogenous Panels." *Journal of Econometrics,* 115: 53-74.

Jorgenson, Dale. 1963. "Capital Theory and Investment Behavior." *American Economic Review*, 53(2): 247-259.

Jorgenson, Dale and Peter Wilcoxen. 1992. "Reducing U.S. Carbon Dioxide Emissions: The Cost of Different Goals." In *Advances in the Economics of Energy and Resources Vol. 7*. ed J.R. Moroney, Greenwich, CT: JAI Press. 195-227.

Judson, Ruth and Anne Owen. 1999. "Estimating Dynamic Panel Data Models: A Guide for Macroeconomists." *Economic Letters,* 65: 9-15.

Kao, Chihwa. 1999. "Spurious Regression and Residual-Based Tests for Cointegration Panel Data." *Journal of Econometrics,* 90: 1-44.

Kao, Chihwa and Min-Hsien Chiang.2000. "On the Estimation and Inference of a Cointegrated Regression in Panel Data." *Advances in Econometrics,* 15: 179-222.

Kaufmann, Robert. 1991. "Oil Production in the Lower 48 States, Reconciling Curve Fitting and Econometric Models." *Resources and Energy,* 31: 111-127.

Kaufmann, Robert. and Cutler Cleveland. 2001. "Oil Production in the Lower 48 States: Economic, Geological. and Institutional Determinants." *Energy Journal,* 22: 27-49.

Kiviet, Jan. 1995. "On Bias, Inconsistency and Efficiency of Various Estimators in Dynamic Panel Data Models." *Journal of Econometrics,* 68: 53-78.

Levin, A. and C. F. Lin. 1992. "Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties." discussion paper #92-93 University of California-San Diego.

Maddala, G.S. and Shaowen Wu. 1999. "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test." *Oxford Bulletin of Economics and Statistics,* 61: 631-652.

Moroney, John 1972. *The Structure of Production in America*. Chapel Hill: The University of North Carolina Press.

Moroney, John, 1997, *Exploration, Development, and Production: Texas Oil and Gas, 1970-1995*, Greenwich CT.: JAI Press.

Moroney, John and M. Douglas Berg. 1999. “An Integrated Model of Oil Production.” *The Energy Journal*, 20: 105-124.

Nickell, Stephen. 1981. “Biases in Dynamic Models with Fixed Effects.” *Econometrica*, 49: 1417-1426.

O hUllachain, Breanden and Richard Matthews. 1996. “Restructuring of Primary Industries: Technology, Labor, and Corporate Strategy and Control in the Arizona Copper Industry.” *Economic Geography*, 72: 196-215.

Pedroni, Peter. 2000. “Fully Modified OLS for Heterogeneous Cointegrated Panels.” *Advances in Econometrics*, 15: 93-130.

Pesaran, Hashem 2003. “A Simple Panel Unit Root Test in the Presence of Cross Section Dependence” *Cambridge Working Papers in Economics 0346*, Faculty of Economics (DAE), University of Cambridge.

Pindyck, Robert 1978. “The Optimal Exploration and Development of Exhaustible Natural Resources” *Journal of Political Economy*, 86(5): 841-861.

Scherer, Frederic. 1980. *Industrial Market Structure and Economic Performance* 2nd edition. Chicago: Rand McNally College Publishing Company.

Schmalansee, Richard, Thomas Stoker, and Ruth Judson. 1998. “World Carbon Dioxide Emissions: 1950-2050.” *Review of Economics and Statistics* 80:2: 15-27.

Stigler, George. 1958. “The Economies of Scale.” *Journal of Law and Economics*, 1: 54-71.

APPENDIX A

BIAS IN DYNAMIC PANEL DATA MODELS AND THE ANDERSON AND HSIAO AND ARELLANO AND BOND PROCEDURES

The simplest method of describing the bias in dynamic panel models is to specify an autoregressive equation.

A1 $$y_{it} = \mu_i + \lambda y_{i,t-1} + \varepsilon_{it}$$

The common procedure used to estimate this model is to use deviations from cross-sectional averages. The result is that A1 is transformed to

A2 $$y_{it} - \frac{1}{T-1}\sum_{t=2}^{T} y_{it} = \lambda(y_{i,t-1} - \frac{1}{T-1}\sum_{t=2}^{T} y_{it-1}) + \varepsilon_{it} - \frac{1}{T-1}\sum_{t=2}^{T}\varepsilon_{it}$$

The disturbance term in A2 is correlated with $y_{i,t-1}$ since $\frac{1}{T-1}\sum_{t=2}^{T}\varepsilon_{it}$ includes ε_{it-1}.

Anderson and Hsiao and Arellano and Bond have proposed methods to correct for the bias in the LSDV estimator. The Anderson and Hsiao approach is a two-stage process in which A1 is estimated in first differences as:

A3 $$\Delta y_{it} = \lambda \Delta y_{i,t-1} + \Delta\varepsilon_{it}$$

$\Delta y_{i,t-1}$ and $\Delta\varepsilon_{it}$ are still correlated since $\Delta y_{i,t-1} = y_{i,t-1} - y_{i,t-2}$ and $\Delta\varepsilon_{it} = \varepsilon_{it} - \varepsilon_{it-1}$. However, $\Delta y_{i,t-2}$ and y_{it-2} are both correlated with $\Delta y_{i,t-1}$ and uncorrelated with $\Delta\varepsilon_{it}$ so that either is a valid instrument for $\Delta y_{i,t-1}$. Most studies prefer to use y_{it-2} because this instrument usually has smaller variance.

The Arellano and Bond approach is to some degree an extension of the Anderson and Hsiao method. The model is estimated in first differences, however; more instruments are used. At least three time-periods of data are necessary to estimate A3 with instrumental variables. In the Arellano and Bond method, the instrument in period 3 is identical to the instrument used by Anderson and Hsiao. However, in period four, y_{i1} and y_{i2} are both valid instruments for $\Delta y_{i,3}$ and in period T-1 the set of valid instruments for $\Delta y_{i,T-1}$ is $(y_{i,T-2}, y_{i,T-3},......, y_{i1})$. These instruments are then used along with moment conditions to estimate A3.3 through GMM procedures.

APPENDIX B

TABLES AND FIGURES FOR CHAPTER III

Table B3.1a. Results of Unit Root Tests for lnProduction

Test Procedure	Constant and Trend (p-value)	Constant No Trend (p-value)	No Constant No Trend (p-value)
Levin Lin	0.9758	0.9526	0.000
IPS	0.3202	1.0000	NA
Maddala Wu (ADF)	0.0466	1.0000	0.0000
Maddala Wu (Phillips-Perron)	0.1614	1.0000	0.0000

All tests have a null hypothesis of a unit root. P-values indicate the minimum critical value at which you can reject the null hypothesis. Values less than .05 indicate a rejection of a unit root at the .05 level.

Table B3.1b. Results of Unit Root Tests for lnPrice

Test Procedure	Constant and Trend (p-value)	Constant No Trend (p-value)	No Constant No Trend (p-value)
Levin Lin	0.0000	0.0000	0.0098
IPS	1.000	0.0000	NA
Maddala Wu (ADF)	1.000	0.0007	0.9204
Maddala Wu (Phillips-Perron)	1.000	0.0018	0.9084

All tests have a null hypothesis of a unit root. P-values indicate the minimum critical value at which you can reject the null hypothesis. Values less than .05 indicate a rejection of a unit root at the .05 level.

Table B3.1c. Results of Unit Root Tests for lnReserves

Test Procedure	Constant and Trend (p-value)	Constant No Trend (p-value)	No Constant No Trend (p-value)
Levin Lin	0.0045	0.0549	0.0000
IPS	.0030	0.6983	NA
Maddala Wu (ADF)	0.0063	0.6859	0.0001
Maddala Wu (Phillips-Perron)	0.0174	0.9648	0.0000

All tests have a null hypothesis of a unit root. P-values indicate the minimum critical value at which you can reject the null hypothesis. Values less than .05 indicate a rejection of a unit root at the .05 level.

Table B3.2. A Comparison of LSDV and Dynamic OLS Estimates

Variable	LSDV estimates (9)	DOLS estimates (9a)
Lnoilprice	.2248 (.0311)	.1488 (.0367)
Lnreserves	.7420 (.0316)	.8024 (.0385)
1986	.1732 (.0382)	.2035 (.0379)
gulf1	.0184 (.0261)	.0718 (.0266)
gulf2	-.3137 (.0369)	-.2619 (.0398)
R-square	0.955	0.958

Standard errors in parentheses

Table B3.3. Estimation Results for Equation 3.12

Variable	LSDV coefficient estimates	Arellano Bond estimates	Anderson Hsiao estimates	Bias Corrected LSDV estimates
lagged lnproduction	0.567 (.039)	.585 (.041)	0.890 (.172)	0.772 (.047)
Lnreserves	0.080 (.025)	.070 (.027)	0.014 (0.037)	.053 (.024)
Lnprice	0.056 (.020)	.055 (.020)	0.060 (.025)	.059 (.019)
1986	0.005 (.0158)	.004 (.016)	-.017 (.017)	-.005 (.016)
gulf1	-0.001 (.011)	-.0002 (.011)	.018 (.015)	-.0016 (.0109)
gulf2	0.011 (.015)	.011 (.015)	-.023 (.023)	.001 (.016)

Standard errors in parentheses.

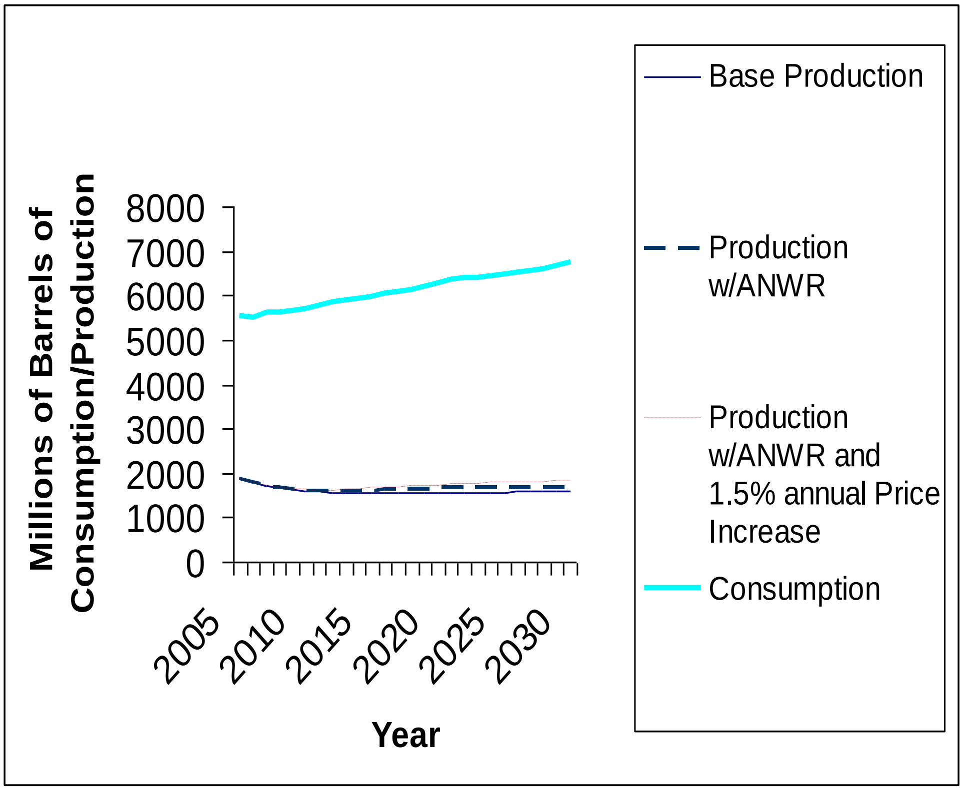

Figure B3.1. Consumption and 3 Production Projections, 2005 – 2030

VITA

Name:	David A. Brightwell
Address:	Illinois Commerce Commission 527 E. Capitol Avenue Springfield, IL. 62701
Email Address:	DavidABrightwell@neo.tamu.edu
Education:	B.S., Political Science, Illinois State University, 1992 M.S., Applied Economics, Illinois State University, 2002 Ph.D. Economics, Texas A&M University, 2008